ISRAEL JOURNAL OF MATHEMATICS **131** (2002), 221–257

PRESSURE AND EQUILIBRIUM STATES FOR COUNTABLE STATE MARKOV SHIFTS

BY

DORIS FIEBIG

Institut für Mathematische Stochastik, Universität Göttingen
Lotzestr. 13, 37083 Göttingen, Germany
e-mail: fiebig@math.uni-goettingen.de

AND

ULF-RAINER FIEBIG

Institut für Angewandte Mathematik, Universtität Heidelberg
Im Neuenheimer Feld 294, 69120 Heidelberg, Germany
e-mail: fiebig@math.uni-heidelberg.de

AND

MICHIKO YURI

Department of Business Administration, Sapporo University
Nishioka, Toyohira-ku, Sapporo 062, Japan
e-mail: yuri@math.sci.hokudai.ac.jp

ABSTRACT

We give a general definition of the topological pressure $P_{top}(f, S)$ for continuous real valued functions $f: X \to \mathbb{R}$ on transitive countable state Markov shifts (X, S). A variational principle holds for functions satisfying a mild distortion property. We introduce a new notion of Z-recurrent functions. Given any such function f, we show a general method how to obtain tight sequences of invariant probability measures supported on periodic points such that a weak accumulation point μ is an equilibrium state for f if and only if $\int f^- d\mu < \infty$. We discuss some conditions that ensure this integrability. As an application we obtain the Gauss measure as a weak limit of measures supported on periodic points.

Received January 15, 2001

0. Introduction

Topological pressure, the variational principle and the existence of equilibrium states for continuous functions on shifts of finite type have been studied by Bowen [B] and Ruelle [R]. This work has been extended to countable state Markov shifts by Walters, Gurevic and Savchenko, and Sarig, [W2], [GS], [S1, S2]. Our work continues and also complements this study. We neither use the Ruelle–Perron–Frobenius operator, nor do we assume Hölder continuity. This allows us to obtain a variational principle and a theory of equilibrium states for a wider class of functions.

We give an outline of the paper and then compare our results with those of Gurevic–Savchenko and Sarig.

Section 1 starts with two natural definitions of topological pressure for continuous functions on two-sided countable state Markov shifts, one via approximations from inside, the other more computational via growth rates of weights of loops. Under the mild distortion condition $D_n(f)/n \to 0$ these two notions agree, which allows one to compute the pressure for such functions via loop-counting methods at any vertex a.

The measure theoretical pressure is introduced in Section 2 and a variational principle is shown for functions f satisfying $D_n(f)/n \to 0$.

Sections 3 and 4 contain the study of equilibrium states. We give a new definition of functions which are Z-recurrent at a vertex a. In the finite entropy case the function $f = 0$ is Z-recurrent if and only if the Markov shift is positive recurrent. Then we assign sequences of measures defined on periodic points that visit a to a function which is Z-recurrent at a. Now Z-recurrence (where we additionally assume $P_{top}(f) < \infty$ and the distortion property $\sup_n D_n(f) < \infty$) ensures that these sequences are tight, thus they have weak accumulation points. In Section 4 we study when such an accumulation point is an equilibrium state.

Section 5 shows that all the results easily carry over to the setting of one-sided Markov shifts.

In Section 6 we briefly discuss the different distortion properties. In particular, we show that Hölder continuous functions satisfy the condition $\sup_n D_n(f) < \infty$, and give an example for a non-Hölder continuous function where our results apply.

In Section 7 we apply our main result about equilibrium states (Theorem 4.2) to the Gauss map.

Finally, Section 8 discusses when it is possible to define the topological pressure by the growth rates of the weights of all periodic points instead of just considering those that visit a fixed vertex.

In his work [S1, S2], Sarig considers Hölder continuous functions on countable state Markov shifts. Hölder continuous functions satisfy the distortion property $\sup_n D_n(f) < \infty$, Observation 6.3, and thus also $D_n(f)/n \to 0$. His notion of topological pressure coincides with ours for Hölder continuous functions, as Theorem 1.9 shows. Sarig proved the Variational Principle for Hölder continuous functions where the associated Ruelle–Perron–Frobenius operator maps the constant function 1 to a bounded function (i.e., $||\mathcal{L}_f 1||_\infty < \infty$). This might be a severe restriction. It implies $\sup_{x\in X} f(x) < \infty$, thus excludes all functions f with $\inf_{x\in X} f(x) > -\infty$ on Markov shifts given by a graph with unbounded in-degree, as the Bernoulli shift or, as an example with finite entropy, the Markov shift given by the graph with vertex set $\mathbb{N}$ and for each $n \in \mathbb{N}$ there is an edge from n to $n+1$ and an edge from n to 1. Since Hölder continuity implies $D_n(f)/n \to 0$, Theorem 2.4 generalizes Sarig's Variational Principle. He showed that an equilibrium state (in a more general sense) exists for positive recurrent functions (and is actually unique). Sarig's results from [S1, S2] imply that functions positive recurrent are also Z-recurrent in the sense of Definition 3.1.

In [GS], Gurevic and Savchenko consider functions which depend only on the zero-coordinate. These are Hölder continuous, but do not have to satisfy the condition $||\mathcal{L}_f 1||_\infty < \infty$. Therefore their results do not follow from Sarig's results. They prove the Variational Principle for bounded functions. Their notion of positive recurrence coincides with that of Sarig, but their notion of equilibrium states is different from Sarig's and from our definition. However, using [GS, Prop. 4.3] one can show that an ergodic good measure (Definition 2.1) is an equilibrium state in the sense of Definition 3.1 iff it is an equilibrium state in the sense of [GS]. They show that a function is positive recurrent iff an equilibrium state exists and in this case the equilibrium state is unique.

The referee has pointed out that Sarig improved the above-mentioned results to the case of functions with $\sup f < \infty$, summable variations, and which not necessarily have to satisfy the condition $||\mathcal{L}_f 1||_\infty < \infty$, [S3]. These new results then also imply the above-cited results of [GS].

We fix some notation. Let E be a countable set. Let $E^{\mathbb{Z}}$ be endowed with the product topology of the discrete topology on E. The left shift map $\sigma\colon E^{\mathbb{Z}} \to E^{\mathbb{Z}}$ is the homeomorphism defined by $(\sigma x)_n := x_{n+1}$, $n \in \mathbb{Z}$ for all $x = (x_n)_{n\in\mathbf{Z}} \in E^{\mathbb{Z}}$. Given a subset X of $E^{\mathbb{Z}}$, a point $x = (x_n)_{n\in\mathbb{Z}} \in X$ and integers $-\infty < n \leq m < \infty$ let $x[n,m]$ and $x[n,m+1)$ denote the block $x_n,\ldots,x_m$ and let ${}_n[x_n,\ldots,x_m]$ denote the cylinder set $\{y \in X | y[n,m] = x[n,m]\}$. The induced topology on X is generated by the cylinder sets. If X is shift invariant, $\sigma(X) = X$, then $(X, \sigma|_X)$ is called a subshift. We denote the points of period n in X by $\mathrm{Per}_n(S) := \{x \in X | S^n x = x\}$.

Let $G = (V, E)$ be a graph with a countable vertex set V and a countable set of directed edges E. We consider only graphs where each vertex has at least one out-going and one in-coming edge. Let $i\colon E \to V$ and $t\colon E \to V$ denote the initial and terminal vertex maps. The **Markov shift** (X, S) defined by G is a subshift of $(E^{\mathbb{Z}}, \sigma)$ where

$$X := \{x = (x_n)_{n\in\mathbb{Z}} \in E^{\mathbb{Z}} | t(x_n) = i(x_{n+1}) \text{ for all } n \in \mathbb{Z}\}.$$

That means X is the set of bi-infinite walks on the directed edges of the graph G, and S is the restricted shift map. We want to emphasize here that we consider all graphs with countable vertex and edge set, and not only graphs where between two vertices there is at most one edge. So we definitely allow parallel edges. In this sense our approach is more general than that of Sarig, and Gurevic and Savchenko.

Note that X is compact iff G is a finite graph iff (X, S) is a shift of finite type (SFT). Transitivity of S means irreducibility of G, that is for every pair of vertices α and β of G there is a path from α to β. Only transitive Markov shifts are considered in this paper. We say that a block $w \in E^k$, $k \geq 1$, is an S-block or a path in S if there is a point $x \in X$ with $x[1, k] = w$. Let $|w|$ denote the length k of the block w.

Remark 0.1: We consider loop counting in different matrix presentations of a Markov shift; later, in Remark 3.4, we shall do the same for the notion of positive recurrence.

Let $G = (V, E)$ be a countable directed graph defining a transitive Markov shift $X = \{x = (x_n)_{n\in\mathbb{Z}} \in E^{\mathbf{Z}} | t(x_n) = i(x_{n+1}) \text{ for all } n \in \mathbb{Z}\}$. Let A be the matrix indexed with the vertex set V and $A_{a,a'} = \#\{e \in E | i(e) = a, t(e) = a'\}$. Note that $A_{a,a'} = \infty$ is allowed.

Now define a 0-1 matrix B, indexed by the edge set E, by $B_{e,e'} = 1$ iff $t(e) = i(e')$. Let

$$X' := \{x' = (x'_n)_{n\in\mathbb{Z}} \in E^{\mathbb{Z}} | B_{x'_n, x'_{n+1}} = 1 \text{ for all } n \in \mathbb{Z}\}.$$

Note that obviously $X' = X$. Let H be the graph defined by the matrix B, i.e., the vertex set of H is $V_H = E$ and for e, $e' \in V_H$ there is an edge from vertex e to vertex e' iff $B_{e,e'} = 1$. Note that H has no parallel edges.

We have that $B^n_{e,e} = \#\{x \in \mathrm{Per}_n(X)|x_0 = e\}$. Thus $B^n_{e,e} \leq A^n_{a,a}$ where $a = i(e)$ is the initial vertex of the edge e in the graph G. Now fix a loop u in G beginning with edge e. Then to a loop w in G of length n at vertex a assign the point $x \in \mathrm{Per}_{n+|u|}(X)$ with $x[0,|u|) = u$, $x[|u|,|u|+n) = w$. This assignment is injective and shows $A^n_{a,a} \leq B^{n+|u|}_{e,e}$. Thus $\limsup_n 1/n \log A^n_{a,a} = \limsup_n 1/n \log B^n_{e,e}$. Since $\limsup_n 1/n \log B^n_{e,e}$ is independent of e, [G], thus also $\limsup_n 1/n \log A^n_{a,a}$ is independent of a. Note that $A_{a,a'} = \infty$ for some a, $a' \in V$ implies $\limsup_n 1/n \log A^n_{a,a} = \infty$.

We recall some basic facts about the pressure for continuous functions on compact subshifts.

Given a space X with a selfmap S and a function $f\colon X \to \mathbb{R}$ let

$$S_n f x := \sum_{i=0}^{n-1} f(S^i x), \quad n \in \mathbb{N},\ x \in X.$$

Suppose that (X,S) is a compact subshift and $f\colon X \to \mathbb{R}$ continuous. Let β be the zero-partition, i.e., the partition into the sets $[i]_0 := \{x \in X|x_0 = i\}$. Let $\beta(n) := \beta \vee S^{-1}\beta \vee S^{-2}\beta \vee \ldots \vee S^{-(n-1)}\beta$. Let $W_n(f,S) = W_n(f) := \sum_{B\in\beta(n)} \sup_{x\in B} \exp(S_n f x)$. The sequence $W_n(f)$ is submultiplicative, thus $P(f) := \lim_{n\to\infty}(1/n)\log W_n(f)$ exists and agrees with $\inf_n(1/n)\log W_n(f)$.

The number $P(f) \in \mathbb{R}$ is the topological pressure of f.

The above definition is equivalent to a more general one via spanning sets [W1], which can be used whenever X is compact metric and S, f are both continuous. But the above suffices for our purposes.

1. Topological pressure

For continuous functions f on transitive Markov shifts we define the inner pressure $P_{in}(f)$ by approximations from inside and the topological pressure $P_{top}(f)$ by counting weights of loops at fixed vertices. These two notions agree if the function satisfies a mild distortion property.

Let (X,S) be a transitive Markov shift given by a countable directed graph $G = (V,E)$ and $f\colon X \to \mathbb{R}$ a continuous function.

Definition 1.1: The inner pressure of f is

$$P_{in}(f,S) := \sup\{P(f|_Y)|(Y,S|_Y) \text{ is a transitive SFT inside } (X,S)\}.$$

One important question is when this quantity can be computed by loop-counting methods, i.e., when $P_{in}(f) = P_{top}(f)$, where the latter is defined as follows. Recall that $i\colon E \to V$ denotes the initial vertex map.

Definition 1.2: For $a \in V$ and $n \in \mathbb{N}$ let $P(n,a) := \{x \in \mathrm{Per}_n(S) | i(x_0) = a\}$. Let

$$Z_n(f,S,a) = \sum_{x \in P(n,a)} \exp(S_n f x)$$

where $Z_n(f,S,a) := 0$ if $P(n,a) = \emptyset$. The topological pressure of f is

$$P_{top}(f,S) := \sup_{a \in V} \limsup_{n \to \infty} \frac{1}{n} \cdot \log Z_n(f,S,a).$$

If S is clear from the context we simply write $P_{in}(f), Z_n(f,a)$, and $P_{top}(f)$, respectively. Using the notation from Remark 0.1, in the case $f = 0$ we have $Z_n(f,S,a) = A^n_{a,a}$ and thus

$$P_{top}(0) = \limsup_n \frac{1}{n} \log A^n_{a,a} = \limsup_n \frac{1}{n} \log B^n_{e,e},$$

by Remark 0.1. Since $\limsup_n 1/n \log B^n_{e,e} = \sup\{h_{top}(Y) | Y \text{ a SFT inside } X'\}$, [G], and $X' = X$, we obtain $P_{top}(0) = P_{in}(0)$. For general continuous functions f we will show $P_{top}(f) = P_{in}(f)$ (and in fact this number can be calculated by considering loops at any vertex a) whenever f satisfies the weak distortion property $D_n(f)/n \to 0$, where the distortions are defined as follows.

Definition 1.3: The nth distortion of f is

$$D_n = D_n(f) = D_n(f,S) := \sup_{x,y \in X, x[0,n) = y[0,n)} |S_n f x - S_n f y| \in [0, \infty].$$

Remark 1.4: In general all the distortions may be infinite. However, if (X,S) is a SFT then f is bounded and thus all $D_n(f)$ are finite. Moreover, $D_n(f)/n \to 0$ since f is uniformly continuous (see Section 6).

The distortions allow one to compare values of the functions $S_n f$ on certain periodic points. This will be used repeatedly, for example to prove weak super-multiplicativity properties of the weights $Z_n(f,a)$ as in Lemma 1.6.

LEMMA 1.5: *Let $n_1, n_2, \ldots, n_k \in \mathbb{N}$. For each tuple $(x^1, \ldots, x^k) \in P(n_1,a) \times P(n_2,a) \times \cdots \times P(n_k,a)$ there is a unique point $\phi(x^1, \ldots, x^k) \in P(n_1 + \cdots + n_k, a)$ with $\phi(x^1, \ldots, x^k)[0, n_1 + \cdots + n_k) = x^1[0,n_1)x^2[0,n_2) \cdots x^k[0,n_k)$ and it is true that*

$$\left| \sum_{i=1}^k S_{n_i} f x^i - S_{n_1 + \cdots + n_k} f \phi(x^1, \ldots, x^k) \right| \le \sum_{i=1}^k D_{n_i}(f).$$

Proof: By the triangle inequality the left hand side is bounded by $\sum_{i=1}^k |S_{n_i} f x^i - S_{n_i} f S^{n_1 + \cdots + n_{i-1}} \phi(x^1, \ldots, x^k)|$. Since

$$(S^{n_1 + \cdots + n_{i-1}} \phi(x^1, \ldots, x^k))[0, n_i) = x^i[0, n_i)$$

the ith term of this sum is bounded by $D_{n_i}(f)$. ∎

LEMMA 1.6: *Let a be a vertex. Then*

$$Z_n(f,a)^k \leq Z_{nk}(f,a) \cdot \exp(k \cdot D_n(f)),$$

for all $k \in \mathbb{N}$ and all $n \in \mathbb{N}$ with $P(n,a) \neq \emptyset$.

Proof: Note that $P(n,a) \neq \emptyset$ ensures $Z_{nk}(f,a) > 0$, thus the lemma holds trivially if $D_n(f) = \infty$ (which can happen only in the non-compact case). Lemma 1.5 applied with $n_i = n$, $1 \leq i \leq k$, shows that

$$\exp\left(\sum_{i=1}^{k} S_n f x^i\right) \leq \exp(S_{kn} f \phi(x^1, \ldots, x^k)) \cdot \exp(k \cdot D_n(f)).$$

Since ϕ is injective the result follows by summation over all $(x^1, \ldots, x^k) \in P(n,a)^k$. ∎

First we consider the SFT case. Counting weights of loops at a fixed vertex gives a formula as well as lower bounds for the pressure.

LEMMA 1.7: *Let (X,S) be a transitive SFT with period p defined by a finite graph G. Let $f\colon X \to \mathbb{R}$ be continuous and a a vertex of G. Then*

$$\text{(a)} \quad P(f) = \lim_{n\to\infty} \frac{1}{np} \log Z_{np}(f,a),$$

$$\text{(b)} \quad P(f) \geq \frac{1}{n} \log Z_n(f,a) - \frac{D_n(f)}{n} \quad \textit{for all } n \in \mathbb{N}.$$

Note that (a) *implies $P(f) = P_{top}(f)$ in the sense of Definition 1.2.*

Proof: (a) Let $\beta(n)$ and $W_n(f)$ be defined as at the end of Section 0. Since $\#(P(n,a) \cap B) \leq 1$ for all $B \in \beta(n)$, we obtain

$$Z_n(f,a) \leq \sum_{B\in\beta(n)} \sup_{x\in B} \exp(S_n f x) = W_n(f)$$

for all n. This shows that $\limsup(1/n)\log Z_n(f,a) \leq P(f)$. On the other hand, there is an N such that for every $B \in \beta(np)$ there is a point $z_B \in P(np+Np,a)$ such that $S^i z_B \in B$ for some $0 \leq i < Np$. Any such assignment is at most Np to 1. For every $x \in B$ we have $|S_{np}fx - S_{np+Np}fz_B| \leq Np \cdot \sup|f| + D_{np}(f)$, thus

$$\begin{aligned} W_{np}(f) &\leq \sum_{B\in\beta(np)} \exp(S_{np+Np} f z_B + Np \cdot \sup|f| + D_{np}(f)) \\ &\leq Np \cdot Z_{np+Np}(f,a) \cdot \exp(Np \cdot \sup|f| + D_{np}(f)). \end{aligned}$$

Thus $P(f) = \lim(1/np) \log W_{np}(f) \le \liminf(1/np) \log Z_{np}(f, a)$, since $D_n(f)/n \to 0$ by Remark 1.4. Altogether we have shown that $\lim(1/np) \log Z_{np}(f, a)$ exists and equals $P(f)$.

(b) This is trivial if $P(n, a)$ is empty. If $P(n, a) \neq \emptyset$ then Lemma 1.6 shows that $(1/n) \log Z_n - D_n/n \le (1/nk) \log Z_{nk}$ for all k. With $k = mp$, $m \to \infty$ the latter approaches $P(f)$ as shown in (a). ∎

Now we return to the general case of countable state Markov shifts.

PROPOSITION 1.8: *Let (X, S) be a transitive Markov shift of period p given by a graph G. Let $f\colon X \to \mathbb{R}$ be continuous and a a vertex of G. Then*

$$\text{(a)} \qquad P_{in}(f) \le \liminf_{n\to\infty} \frac{1}{np} \log Z_{np}(f, a),$$

$$\text{(b)} \qquad P_{in}(f) + \frac{D_n(f)}{n} \ge \frac{1}{n} \log Z_n(f, a) \text{ for all } n.$$

Note that (a) *implies $P_{in}(f) \le P_{top}(f)$ in the sense of Definition 1.2.*

Proof: (a) Let $\varepsilon > 0$ and choose a SFT $Y \subset X$ with period p which is given by a subgraph of G that contains a and such that $P(f|_Y) \ge P_{in}(f) - \varepsilon$ if $P_{in}(f) < \infty$ and such that $P(f|_Y) \ge 1/\varepsilon$ if $P_{in}(f) = \infty$. Then, by Lemma 1.7a, $P(f|_Y) = \lim(1/np) \log Z_{np}(f|_Y, a) \le \liminf(1/np) \log Z_{np}(f, a)$, the latter since $Z_{np}(f|_Y, a) \le Z_{np}(f, a)$ for all n. With $\varepsilon \to 0$ the result follows.

(b) We may assume that $D_n(f) < \infty$ and $P(n, a) \neq \emptyset$, since otherwise the inequality is trivial. First assume that $Z_n(f, a) < \infty$. Given $0 < \gamma < 1$ choose a SFT $Y \subset X$ given by a large subgraph of G that contains a and satisfies $Z_n(f|_Y, a) \ge \gamma \cdot Z_n(f, a)$. By Lemma 1.7b

$$P_{in}(f) + \frac{D_n(f)}{n} \ge P(f|_Y) + \frac{D_n(f|_Y)}{n} \ge \frac{1}{n} \log Z_n(f|_Y, a) \ge \frac{\log\gamma}{n} + \frac{1}{n} \log Z_n(f, a).$$

The result follows with $\gamma \to 1$.

If $Z_n(f, a) = \infty$ then choose Y such that $Z_n(f|_Y, a) \ge 1/\gamma$. Lemma 1.7 implies that $P_{in}(f) + D_n(f)/n \ge -(1/n) \log \gamma$; the result follows with $\gamma \to 0$. ∎

The main result of this section shows that the pressure of f can be computed by counting the weights of loops at any fixed vertex whenever f satisfies a weak distortion property.

THEOREM 1.9: *Let (X, S) be a transitive Markov shift of period p given by a graph G. Let $f\colon X \to \mathbb{R}$ be continuous with $D_n(f)/n \to 0$. Then*

$$P_{top}(f) = P_{in}(f) = \lim_{n\to\infty} \frac{1}{np} \cdot \log Z_{np}(f, a) = \limsup_{n\to\infty} \frac{1}{n} \log Z_n(f, a)$$

for every vertex a of G.

Proof: If $P_{in}(f) = \infty$ then Proposition 1.8a shows that

$$\infty = \lim_{n\to\infty} 1/np \cdot \log Z_{np}(f,a).$$

Now suppose that $P_{in}(f) < \infty$. Then by Proposition 1.8 and $D_n(f)/n \to 0$ we obtain

$$\begin{aligned} P_{in}(f) &\le \liminf_{n\to\infty} \frac{1}{np} \log Z_{np}(f,a) \\ &\le \limsup_{n\to\infty} \frac{1}{np} \log Z_{np}(f,a) \le P_{in}(f). \end{aligned}$$

Thus $P_{in}(f) = \lim_{n\to\infty}(1/np) \cdot \log Z_{np}(f,a) = \limsup_{n\to\infty}(1/n)\log Z_n(f,a)$ for all a, the latter since $Z_n(f,a) = 0$ if n is not a multiple of p. By definition this implies $P_{in}(f) = P_{top}(f)$. ∎

Remark 1.10: Given a vertex a and n, $m \in \mathbb{N}$ such that $P(n,a) \neq \emptyset$, $P(m,a) \neq \emptyset$, then

$$Z_n(f,a) \cdot Z_m(f,a) \le Z_{n+m}(f,a) \cdot \exp(D_n(f) + D_m(f)).$$

The proof of Lemma 3.7 in Section 3 shows the argument. In the case $D_n(f)/n \to 0$ this weak supermultiplicativity can be used to give a direct proof that $1/np \cdot \log Z_{np}(f,a)$ is a convergent sequence (where p is the period of the Markov shift). However, the above approximation arguments are still needed to identify the limit.

2. The Variational Principle

In this section we state and prove the Variational Principle (Theorem 2.4) for continuous functions f with $D_n(f)/n \to 0$. This extends results of [G], [GS], [S1]. At the end of this section we supply a variation of the proof, more in the spirit of P. Walters' proof for compact spaces [W1, Thm. 8.6+Thm. 9.10].

Definition 2.1: Let (X,S) be a transitive Markov shift and $f\colon X \to \mathbb{R}$ a continuous function. A shift invariant Borel probability measure μ on X is a good measure for f if $\int f^- d\mu < \infty$. The measure theoretical pressure of f is

$$P_{measure}(f,S) := \sup\{h_\mu(S) + \int f d\mu | \mu \text{ is a good measure for } f\}.$$

Remark 2.2: The quantity $h_\mu(S)+\int f d\mu$ would still make sense if $\int f^- d\mu = \infty$ but $h_\mu(S)+\int f^+ d\mu < \infty$. However, in this case $h_\mu(S)+\int f d\mu = -\infty$, i.e., such measures do not contribute to $P_{measure}(f,S)$.

Remark 2.3: The standard Variational Principle for compact spaces implies immediately that $P_{in}(f) \le P_{measure}(f)$.

THEOREM 2.4 (Variational Principle): *Let (X,S) be a transitive Markov shift and let $f\colon X \to \mathbb{R}$ be a continuous function with $D_n(f)/n \to 0$. Then*

$$P_{top}(f) = P_{in}(f) = P_{measure}(f).$$

Proof: We give a proof in the mixing case. The general case, where S has a period $p \in \mathbb{N}$, follows by considering arithmetic progressions $n = mp$. By Theorem 1.9 and Remark 2.3 we have $P_{top}(f) = P_{in}(f) \le P_{measure}(f)$, thus it remains to show that $P_{measure}(f) \le P_{top}(f)$. We may assume that $P_{top}(f) < \infty$. Let μ be a good measure for f.

For $N \in \mathbb{N}$ let $f_N\colon X \to \mathbb{R}$ be defined by $f_N(x) := \min(f(x), N)$. Since $\int f^- d\mu < \infty$ we have $\int f_N d\mu \to \int f d\mu$ by monotone convergence, thus it suffices to show that

$$h_\mu(S) + \int f_N d\mu \le P_{top}(f) \quad \text{for all } N \in \mathbb{N}.$$

Fix $N \in \mathbb{N}$. Identify the set E of edges with $\mathbb{N}$; this induces an ordering on E. We first choose a suitable generating sequence $\alpha_k, k \in \mathbb{N}$, of finite partitions of X. Fix a sequence $(r_k)_{k\in\mathbb{N}}$ of integers with $r_k > k$ for all k and such that $\mu(\{x \in X | x_0 > r_k\}) \cdot \log(k+2) \to 0$. Let α_k be the partition of X into the sets

$$\begin{aligned} [a]_0 &:= \{x \in X | x_0 = a\} \quad \text{for } a \le k, \\ D_k &:= \{x \in X | k < x_0 \le r_k\} \quad \text{and} \\ C_k &:= \{x \in X | x_0 > r_k\}. \end{aligned}$$

Thus α_k is a partition with $k+2$ atoms and, by the choice of the sequence r_k, we have

$$\lim_{k\to\infty} \mu(C_k) \cdot \log(\#\alpha_k) = 0.$$

Since the sequence $\alpha_k, k \in \mathbb{N}$ is generating, $h_\mu(S) = \lim_{k\to\infty} h_\mu(\alpha_k, S)$. Now fix k for the moment and let $\alpha := \alpha_k$. Let $\alpha(n) := \alpha \vee S^{-1}\alpha \vee S^{-2}\alpha \vee \cdots \vee S^{-(n-1)}\alpha$. For $P \in \alpha(n)$ let $g(n,P) = \sup\{S_n f_N x | x \in P\}$. Then $g(n,P) < \infty$ since $\sup f_N \le N < \infty$. For $A, B \in \alpha$ let

$$\alpha(n,A,B) = \{P \in \alpha(n) | P \text{ is contained in } A \cap S^{-(n-1)}B\}.$$

By S-invariance of μ we obtain

$$\begin{aligned}\int f_N d\mu &= \frac{1}{n}\int S_n f_N d\mu \\ &\le \frac{1}{n}\sum_{A,B\in\alpha}\sum_{P\in\alpha(n,A,B)} g(n,P)\cdot\mu(P) \\ &= \frac{1}{n}\sum_{A,B\in\alpha}\mu(A\cap S^{-(n-1)}B)\cdot\sum_{P\in\alpha(n,A,B)} g(n,P)\cdot\mu(P|A\cap S^{-(n-1)}B).\end{aligned}$$

Note that $0\le H_\mu(\alpha\vee S^{-(n-1)}\alpha)\le 2\log(\#\alpha)$ implies

$$\lim_{n\to\infty}\frac{1}{n}H_\mu(\alpha(n)) = \lim_{n\to\infty}\frac{1}{n}H_\mu(\alpha(n)|\alpha\vee S^{-(n-1)}\alpha).$$

This gives

$$\begin{aligned}&h_\mu(\alpha_k,S)+\int f_N d\mu \\ &= \lim_{n\to\infty}\frac{1}{n}[H_\mu(\alpha(n))+\int S_n f_N d\mu] \\ &\le \limsup_{n\to\infty}\frac{1}{n}\sum_{A,B\in\alpha}\mu(A\cap S^{-(n-1)}B)\cdot\sum_{P\in\alpha(n,A,B)}\mu(P|A\cap S^{-(n-1)}B) \\ &\qquad\qquad\cdot[-\log\mu(P|A\cap S^{-(n-1)}B)+g(n,P)] \\ &\le \limsup_{n\to\infty}\frac{1}{n}\sum_{A,B\in\alpha}\mu(A\cap S^{-(n-1)}B)\cdot\log\Bigg(\sum_{P\in\alpha(n,A,B)}\exp(g(n,P))\Bigg).\end{aligned}$$

The last estimate holds by [W1, Lemma 9.9], which states that given real numbers $a_1,\ldots,a_k$ and $p_i\ge 0$ with $\sum_{i=1}^k p_i = 1$, then $\sum_{i=1}^k p_i(a_i-\log p_i)\le \log(\sum_{i=1}^k\exp(a_i))$.

We estimate the sequences $(1/n)\log(\sum_{P\in\alpha(n,A,B)}\exp(g(n,P)))$.

First consider the case that $A=C_k$ or $B=C_k$. The rough estimate $f_N\le N$ gives

$$\begin{aligned}\frac{1}{n}\log\Bigg(\sum_{P\in\alpha(n,A,B)}\exp(g(n,P))\Bigg) &\le \frac{1}{n}\log(\sum_{P\in\alpha(n,A,B)}\exp(n\cdot N)) \\ &= \frac{1}{n}\log(\#\alpha(n,A,B)\cdot\exp(n\cdot N)) \\ &\le \log(\#\alpha)+N.\end{aligned}$$

Thus

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}&\sum_{A,B\in\alpha, A=C_k \text{ or } B=C_k}\mu(A\cap S^{-(n-1)}B)\cdot\log\Bigg(\sum_{P\in\alpha(n,A,B)}\exp(g(n,P))\Bigg) \\ &\le 2\mu(C_k)\cdot(\log(\#\alpha_k)+N).\end{aligned}$$

In the case that $A \neq C_k$ and $B \neq C_k$, let d be a fixed vertex of the graph G. By Theorem 1.9 we know $P_{top}(f) = \lim_n (1/n) \log Z_n(f,d)$.

Fix M so that $D_n(f) < \infty$ for all $n \geq M$. For every edge $a \leq r_k$ let $w(a)$ be a path starting in vertex d and let $u(a)$ be a path ending in vertex d such that $w(a)au(a)$ is a path. Since S is mixing, we may assume that for some $m > M$ these paths have length $m = |w(a)| = |u(a)|$ for all $a \leq r_k$. Since $D_m(f) < \infty$, there is some M_k such that $|S_m fx| \leq M_k$ whenever $x[0,m) = w(a)$ or $x[0,m) = u(a)$ for some $a \leq r_k$.

Now consider $A, B \in \alpha$ with $A \neq C_k$ and $B \neq C_k$. Let $\varepsilon > 0$ and $n \geq M$. For each $P \in \alpha(n,A,B)$ choose a point $x^P \in P$ with $g(n,P) \leq S_n f_N x^P + \varepsilon$. Then in particular $g(n,P) \leq S_n f x^P + \varepsilon$. Let $w^P = x^P[0,n), a = x_0^P$ and $b = x_{n-1}^P$. Now $x^P \in P$ implies $a, b \leq r_k$ and thus $w(a)w^P u(b)$ is a loop of length $n + 2m$ at vertex d. Let $y^P \in P(n+2m, d)$ with $y^p[0, n+2m) = w(a)w^P u(b)$. Since $(S^m y^P)[0,n) = x^P[0,n)$ we get $S_n f(x^P) \leq S_n f(S^m y^P) + D_n(f)$. Since $y^P[0,m) = w(a)$ we get $S_m f(y^P) \geq -M_k$ and since $(S^{n+m}y^P)[0,m) = u(b)$ we get $S_m f(S^{n+m} y^P) \geq -M_k$. Thus $\exp(S_n f x^p) \leq \exp(S_{n+2m} f y^p + D_n(f) + 2M_k)$. Thus

$$\begin{aligned}
\sum_{P\in\alpha(n,A,B)} \exp g(n,P) &\leq \sum_{P\in\alpha(n,A,B)} \exp(S_n f(x^P) + \varepsilon) \\
&\leq \sum_{P\in\alpha(n,A,B)} \exp(S_{n+2m} f(y^P) + D_n(f) + 2M_k + \varepsilon) \\
&= \exp(D_n(f) + 2M_k + \varepsilon) \cdot \sum_{P\in\alpha(n,A,B)} \exp(S_{n+2m} f y^p) \\
&\leq \exp(D_n(f) + 2M_k + \varepsilon) Z_{n+2m}(f,d)
\end{aligned}$$

and $(D_n(f) + 2M_k + \varepsilon)/n \to 0$ implies

$$\limsup_n \frac{1}{n} \log\left(\sum_{P\in\alpha(n,A,B)} \exp g(n,P)\right) \leq \lim_n \frac{1}{n} \log Z_{n+2m}(f,d) = P_{top}(f),$$

the latter by Theorem 1.9. Combining the above estimates gives

$$\begin{aligned}
&h_\mu(\alpha_k, S) + \int f_N d\mu \\
&\leq \limsup_{n\to\infty} \frac{1}{n} \sum_{A,B\in\alpha} \mu(A \cap S^{-(n-1)}B) \cdot \log\left(\sum_{P\in\alpha(n,A,B)} \exp g(n,P)\right) \\
&\leq \sum_{A,B\in\alpha_k - \{C_k\}} \mu(A\cap S^{-(n-1)}B) \cdot P_{top}(f) + 2\mu(C_k) \cdot (\log(\#\alpha_k) + N) \\
&\leq P_{top}(f) + 2\mu(C_k) \cdot (\log(\#\alpha_k) + N).
\end{aligned}$$

Note that $\lim_{k\to\infty} 2\mu(C_k)\cdot(\log(\#\alpha_k)+N) = 0$ by construction, thus with $k \to \infty$ the above yields $h_\mu(S) + \int f_N d\mu \le P_{top}(f)$ and the Variational Principle is proved. ∎

Remark 2.5: We outline an alternative proof, more in the spirit of P. Walters' proof of the Variational Principle for compact spaces. Now the distortions will enter the picture just at the very last step. It makes one believe that the condition $D_n(f)/n \to 0$ is quite natural for a Variational Principle to hold; on the other hand, it makes it hard to speculate about possible weakenings of this condition.

By Theorem 1.9 and Remark 2.3 it suffices to show $P_{measure}(f,S) \le P_{in}(f,S)$.

1. Prove this for functions f which depend only on the zero coordinate. This can be done as in the proof of Theorem 2.4, but the estimates simplify since all the $D_n(f)$ are zero.

2. Show $P_{measure}(f,S) \le P_{in}(f,S) + D_1(f)$ for any continuous function f. If $D_1(f) = \infty$ this is trivial. If $D_1(f) < \infty$ then $h(x) := \sup\{fy | y_0 = x_0\}$ defines a function $h: X \to \mathbb{R}$ which depends only on the zero coordinate and satisfies $f \le h \le f + D_1(f)$. Thus $P_{in}(h,S) \le P_{in}(f,S) + D_1(f)$, [W1, Thm. 9.7] and an S-invariant measure is good for f iff it is good for h. Thus, by $f \le h$ and step 1 we get

$$\begin{aligned} P_{measure}(f,S) &\le P_{measure}(h,S) \\ &= P_{in}(h,S) \\ &\le P_{in}(f,S) + D_1(f). \end{aligned}$$

3. Now apply the above to the shift (X, S^n) endowed with the function $S_n f$, and link the occurring quantities to those for S and f (this is done in Lemma 2.7 at the end of this section). Thus the proof proceeds as follows:

$$\begin{aligned} P_{measure}(f,S) &\le \frac{1}{n} P_{measure}(S_n f, S^n) \text{ by Lemma 2.7a} \\ &\le \frac{1}{n} P_{in}(S_n f, S^n) + \frac{1}{n} D_1(S_n f, S^n) \text{ by step 2} \\ &\le P_{in}(f,S) + \frac{1}{n} D_n(f,S) \end{aligned}$$

by Lemma 2.7b and since $D_n(f,S) = D_1(S_n f, S^n)$. Now $D_n(f,S)/n \to 0$ implies $P_{measure}(f,S) \le P_{in}(f,S)$, which finishes the outline of the alternative proof.

Remark 2.6: One might wonder if Step 1 is a simple consequence of Gurevic and Savchenko's Variational Principle for bounded functions which depend only on the zero coordinate [GS]. This is not so, since $P_{top}(f) = \infty$ for bounded functions

f on a Markov shift with infinite Gurevic entropy. Thus one cannot approximate a function $f \leq 0$ with $\inf f = -\infty$ and $P_{top}(f) < \infty$ by bounded functions and deduce the result. Also Sarig's result is not strong enough, because he uses the condition $||\mathcal{L}_f 1||_\infty < \infty$. This implies $\sup_{x\in X} f(x) < \infty$ and excludes all functions f with $\inf_{x\in X} f(x) > -\infty$ on a Markov shift given by a graph with unbounded in-degree.

Finally we state and prove the lemma used in the last step above.

LEMMA 2.7:

(a) $n \cdot P_{measure}(f, S) \leq P_{measure}(S_n f, S^n)$,

(b) $n \cdot P_{in}(f, S) \geq P_{in}(S_n f, S^n)$.

Proof: (a) Let μ be a good measure for f on the Markov shift (X, S). Then μ is also S^n-invariant. Obviously $S_n f^+ \geq (S_n f)^+$. Replace f by $-f$; now $(-f)^+ = f^-$ yields $S_n(f^-) \geq (-S_n f)^+ = (S_n f)^-$. Thus $\int (S_n f)^- d\mu \leq \int S_n(f^-) d\mu = n \cdot \int f^- d\mu < \infty$ and μ is a good measure for $S_n f$ on the Markov shift (X, S^n). Since $f \geq -f^-$ and $\int f^- d\mu < \infty$ we get $n \cdot \int f d\mu = \int S_n f d\mu$. This and $n \cdot h_\mu(S) = h_\mu(S^n)$ show that $n \cdot (h_\mu(S) + \int f d\mu) \leq P_{measure}(S_n f, S^n)$.

(b) First note that

$$P_{in}(f, S) = \sup\{P(f|_Y, S|_Y) | Y \subset X, Y \text{ compact and } SY = S\}.$$

Now let $Y \subset X$ be compact with $S^n Y = Y$. Let $Z = Y \cup SY \cup \cdots \cup S^{n-1}Y$. Then $SZ = Z$ and Z is compact. Thus $P_{top}(S_n f|_Y, S^n|_Y) \leq P_{top}(S_n f|_Z, S^n|_Z)$ $= n \cdot P_{top}(f|_Z, S|_Z)$, [W1, Thm. 9.8]. By the first remark, which holds for $S_n f, S^n$ as well, the result follows. ∎

3. Z-recurrence

Suppose f is a continuous function on a transitive Markov shift. We define and discuss the notion of Z-recurrence of f at a vertex a. In the case $f = 0, P_{top}(f) < \infty$ this turns out to be equivalent to the notion of positive recurrent Markov shifts. Then we assign sequences of measures to a function that is Z-recurrent at a vertex a and satisfies $P_{top}(f) < \infty$ and $\sup_n D_n(f) < \infty$. These measures will be supported on periodic points that visit a and Z-recurrence of f ensures the existence of weak accumulation points (any such sequence is shown to be tight). In the following section we give conditions which imply that such an accumulation point has to be an equilibrium state.

Throughout this section (X, S) is a transitive Markov shift given by a countable directed graph $G = (V, E)$ with initial vertex map $i: E \to V$, and $f: X \to \mathbb{R}$ is a continuous function.

For any vertex a and $n \geq 1$ we defin $P(n, a) = \{x \in \mathrm{Per}_n(S) | i(x_0) = a\}$ and $Z_n(f, a) = \sum_{x \in P(n,a)} \exp(S_n f x)$. Now let

$$P^*(n, a) = \{x \in P(n, a) | i(x_k) \neq a \text{ for } 1 \leq k \leq n-1\} \quad \text{and}$$
$$Z_n^*(f, a) = \sum_{x \in P*(n,a)} \exp(S_n f x).$$

Definition 3.1: The continuous function f is Z-recurrent at a vertex a if $Z_n(f, a) < \infty$ for all n and

$$\sum_{n=1}^{\infty} n \frac{Z_n^*(f, a)}{Z_n(f, a)} < \infty,$$

where $Z_n^*(f, a)/Z_n(f, a) := 0$ if $Z_n(f, a) = 0$ (i.e., if $P(n, a) = \emptyset$).

Observation 3.2: Later we will mainly consider functions f with $P_{top}(f) < \infty$ and $D_n(f) < \infty$ for all n. In this case $Z_n(f, a) < \infty$ for all n and all $a \in V$, which can be seen as follows. Fix n. The condition $P_{top}(f) < \infty$ implies that there is a k with $Z_{nk}(f, a) < \infty$. If $P(n, a) = \emptyset$ then $Z_n(f, a) = 0$ by definition, and if $P(n, a) \neq \emptyset$ then $Z_n(f, a) < \infty$ by Lemma 1.6.

Z-recurrence holds trivially in some important special cases. For example, suppose that the lengths of the first return loops at a are bounded (think of the Bernoulli shift given by a graph with a single vertex a). Then any function f with $\sup_n D_n(f) < \infty$ and $P_{top}(f) < \infty$ is Z-recurrent at a just because the $Z_n^*(f, a)$ vanish eventually.

QUESTION 3.3: *Suppose that f is Z-recurrent at some vertex. Is f Z-recurrent at every vertex?*

Remark 3.4: We indicate that this holds for the function $f = 0$ under the hypothesis $P_{top}(f) < \infty$ since in this case Definition 3.1 coincides with the definition of positive recurrent Markov shifts. But first we give an argument that positive recurrence does not depend on the matrix presentation chosen for the Markov shift. We give the argument for the mixing case; the general transitive case then follows by using [K, Lemma 7.1.36].

We use the notation from Remark 0.1. Then $Z_n = Z_n(0, a) = A_{a,a}^n$ and thus $P_{top}(0) = \limsup(1/n) \log Z_n = \limsup(1/n) \log A_{a,a}^n = \limsup(1/n) \log B_{e,e}^n$.

Let $\lambda > 0$ so that $P_{top}(0) = \log\lambda$. We show that A is positive recurrent iff B is positive recurrent. For that, define the $V \times V_H$ matrix R with entries 0 or 1 by $R_{a,e} = 1$ iff $i(e) = a$ in the graph G. Define the $V_H \times V$ matrix S with entries 0 or 1 by $S_{e,a} = 1$ iff $t(e) = a$. Then

$$(RS)_{a,a'} = \sum_{e \in V_H} R_{a,e}S_{e,a'} = \#\{e \in V_H | i(e) = a, t(e) = a'\} = A_{a,a'},$$
$$(SR)_{e,e'} = \sum_{a \in V} S_{e,a}R_{a,e'} = R_{t(e),e'} = B_{e,e'}.$$

Thus we have shown $RS = A$ and $SR = B$. Suppose that B is positive recurrent. Then, by [K, Thm. 7.1.3 (d)], $\lambda < \infty$ and there are vectors l, r indexed by V_H with $l, r > 0, l \cdot r < \infty$ and $Br = \lambda r, lB = \lambda l$. Since $r > 0$ and R is a 0-1 matrix with no rows zero, we obtain $Rr > 0$ and $A(Rr) = RSRr = RBr = R\lambda r = \lambda(Rr)$, and since $l > 0$ and S is a 0-1 matrix with no columns zero we obtain $lS > 0$ and $(lS)A = lSRS = lBS = \lambda(lS)$. Furthermore $(lS) \cdot (Rr) = lSRr = lBr = \lambda \cdot lr < \infty$. Now [K, Lemma 7.1.16] shows that A is recurrent, and thus [K, Thm. 7.1.3 (d)] implies that A is positive recurrent. The argument is symmetric in A and B, since it does not use the fact that B is a 0-1 matrix. Thus we have shown A is positive recurrent iff B is positive recurrent.

Now we show that A is positive recurrent iff $f = 0$ is Z-recurrent at any vertex $a \in V$. If A is positive recurrent, then [K, Thm. 7.1.3 (f)] shows that for some $c > 0$ and all large n we have $Z_n(0, a) = A^n_{a,a} \geq c \cdot \lambda^n$. Thus, for N large, $\sum_{n=N}^{\infty} n \cdot Z_n^*/Z_n \leq \sum_{n=N}^{\infty} n/c \cdot Z_n^* \lambda^{-n} < \infty$ and, since all $Z_n(0, a)$ are finite, the function $f = 0$ is Z-recurrent at a.

We give an argument for the converse (however, there should be a more elementary one). If $\sum_{n=1}^{\infty} n \cdot Z_n^*/Z_n < \infty$ at some vertex a, then Theorem 4.2 shows the existence of an equilibrium state for $f = 0$, which is now a measure of maximal entropy for the shift S. Thus since $X' = X, B$ is positive recurrent [G, see also K, Prop. 7.2.13] and thus, as seen above, A is positive recurrent. ∎

QUESTION 3.5: *Is there an elementary proof that in the finite entropy case the function $f = 0$ is Z-recurrent iff the Markov shift is positive recurrent?*

Under the assumption $Z_n(f, a) < \infty$ for all n we are going to define sequences of invariant probability measures supported on periodic points. These periodic points are built from suitable subsets of first return loops at a. It will often be possible to consider the set of all first return loops (Theorem 4.2.d, the Gauss map example in Section 7). However, considering subsets of first return loops will have certain advantages in future applications. For example, it shows that the statement of Theorem 4.2.b is non-void, since the assumptions of this theorem can always be satisfied with suitable finite subsets L_n of first return loops and then the associated measures are trivially good measures. That is the reason why we deal with the extra complication of considering subsets of first return loops.

We first describe the abstract setup.

Let $x \in P^*(k,a)$. Then we say that $x[0,k)$ is a first return loop at a of length k. Let L be an arbitrary subset of the set of all first return loops at a. Define

$$P(n,a,L) := \{x \in P(n,a) | x[0,n) = w_1 \cdots w_r, w_i \in L, 1 \le i \le r\} \quad \text{and}$$
$$Z_{n,L}(f,a) := \sum_{x \in P(n,a,L)} \exp(S_n f x).$$

Whenever a function f and a vertex a have been fixed, then we also write $Z_{n,L}, Z_n$, and Z_n^* where these are meant to be functions of f and a as defined above.

Let δ_x denote the probability measure with $\delta_x(A) = 1$ iff $x \in A$, for every Borel set A.

Definition 3.6: Suppose f is a function with $Z_n(f,a) < \infty$ for all n and L a set of first return loops at vertex a. Then, for every $n \in \mathbb{N}$ with $P(n,a,L) \neq \emptyset$, we define probability measures $\nu_n = \nu_{n,L}$ and $\mu_n = \mu_{n,L}$ by

$$\nu_n := \frac{1}{Z_{n,L}} \cdot \sum_{x \in P(n,a,L)} \exp(S_n f x) \cdot \delta_x \quad \text{and}$$
$$\mu_n := \frac{1}{n} \sum_{i=0}^{n-1} S^i \nu_n.$$

Note that $S^n \nu_n(A) = \nu_n(A)$ for every Borel set A and thus $S\mu_n = \mu_n$, i.e., μ_n is shift invariant.

We supply a little arithmetic for the $Z_{n,L}$ which will be used to prove the next lemma.

LEMMA 3.7: *Suppose that $Z_n(f,a) < \infty$ for all n and $C := \exp(\sup_n D_n(f)) < \infty$. Let L be a subset of the first return loops at a. Then*

$$Z_{n,L} \cdot Z_{k,L} \le C^2 \cdot Z_{n+k,L} \quad \textit{for all } n,k.$$

Proof: Suppose $P(n,a,L) \neq \emptyset$ and $P(k,a,L) \neq \emptyset$ (otherwise the lemma holds trivially). Then Lemma 1.5 applied with $n_1 = n, n_2 = k$ shows $\exp(\sum_{i=1}^{2} S_n f x^i)$ $\le \exp(S_{n+k} f \phi(x^1, x^2)) \cdot C^2$. Since ϕ restricted to $P(n,a,L) \times P(k,a,L)$ is an injective map into $P(n+k,a,L)$, the result follows by summation over all $(x^1,x^2)\in P(n,a,L)\times P(k,a,L)$. ∎

The following estimate for the ν_n-measures of certain cylinder sets will be used repeatedly.

LEMMA 3.8: *Suppose that $Z_n(f,a)<\infty$ for all n and $C:=\exp(\sup_n D_n(f))<\infty$. Let $k\le n$. Suppose that L is a set of first return loops at a with $P(n,a,L)\neq\emptyset, P(k,a,L)\neq\emptyset$, and $Z_{k,L}\ge\gamma Z_k$ for some $\gamma>0$. Let $\nu_n=\nu_{n,L}$ be the measure from Definition 3.6. Then for any set $A=\{x|x[0,k)\in K\}$ where K denotes some subset of the loops of length k at a we have the estimate*

$$\begin{aligned}\nu_n(S^{-j}A)&=\nu_{n,L}(S^{-j}A)\\ &\le\gamma^{-1}C^4\cdot Z_k^{-1}\cdot\sum_{x\in A\cap P(k,a)}\exp(S_kfx)\quad \text{for all } 0\le j<n.\end{aligned}$$

In particular, the same estimate holds for $\mu_n(A)$.

Remark 3.9: With L' denoting the set of all first return loops at a the above estimate becomes $\nu_{n,L}(S^{-j}A)\le\gamma^{-1}C^4\cdot\nu_{k,L'}(A)$ for all $0\le j<n$. If additionally $L=L'$, then one can obviously choose $\gamma=1$.

Proof: In the case $k=n$ the inequality is easy to check (or proceed as in the following proof but omit the z-coordinate). Thus assume $k<n$. If $x\in S^{-j}A$ contributes to $\nu_n(S^{-j}A)$ then $x\in S^{-j}A\cap P(n,a,L)$. Define a map ϑ: $S^{-j}A\cap P(n,a,L)\to(A\cap P(k,a))\times P(n-k,a,L)$ by $\vartheta(x)=(y,z)$, where $y\in P(k,a)$ with $y[0,k)=(S^jx)[0,k)$ and $z\in P(n-k,a,L)$ with $z[0,n-k)=(S^jx)[k,n)$. Then for $(y,z)=\vartheta(x)$ we have that $S_nfx=S_nf(S^jx)\le S_kfy+D_k+S_{n-k}fz+D_{n-k}$ and thus $\exp(S_nfx)\le C^2\cdot\exp(S_kfy)\exp(S_{n-k}fz)$. Since ϑ is injective, summation over all $x\in S^{-j}A\cap P(n,a,L)$ yields

$$Z_{n,L}\cdot\nu_n(S^{-j}A)\le C^2\cdot\left(\sum_{y\in A\cap P(k,a)}\exp(S_kfy)\right)\cdot Z_{n-k,L}.$$

By Lemma 3.7, $Z_{n-k,L}/Z_{n,L}\le C^2/Z_{k,L}$ and the result follows since $Z_{k,L}\ge\gamma Z_k$ by assumption. ∎

If for each n a large enough set L_n has been chosen, and if f is Z-recurrent at a, then the sequence of measures μ_{n,L_n} will have good properties.

THEOREM 3.10: *Suppose (X,S) is a mixing Markov shift given by a graph $G=(V,E)$ where $E=\mathbb{N}$ or $E=\{1,\ldots,N_E\}$ for some $N_E\in\mathbb{N}$. Let $f\colon X\to\mathbb{R}$ be a* *continuous function which is Z-recurrent at vertex a and satisfies $P_{top}(f) < \infty$ and $\sup_n D_n(f) < \infty$. Let $\gamma > 0$.*

Assume that for each n with $P(n,a) \neq \emptyset$ a set of first return loops L_n has been chosen such that

$$Z_{m,L_n} \geq \gamma Z_m \quad \text{for all } 1 \leq m \leq n.$$

Then $\mu_n := \mu_{n,L_n}$ (as defined above) yields a sequence (indexed by $\{n|P(n,a) \neq \emptyset\}$) of invariant probability measures such that

(a) *for every $\varepsilon > 0$ there is some N such that $\mu_n(x_0 > N) < \varepsilon$ for all n,*
(b) *the sequence μ_n is tight,*
(c) *for every $\varepsilon > 0$ there is some N such that $\mu_n\{x|i(x_s) \neq a \text{ for } 0 \leq s \leq N\} < \varepsilon$ for all n.*

Proof: Let $C := \exp(D)$ where $D = \sup_n D_n(f,S)$. (a) Let $\varepsilon > 0$. For $N, l \in \mathbb{N}$ define $B_l(N) = \{x|x[0,l) \text{ is a first return loop at } a \text{ and } x_i > N \text{ for some } 0 \leq i < l\}$ and $A_l(N) = \sum_{x \in B_l(N) \cap P(l,a)} \exp(S_l f x)$. Since $\sum_{l=1}^{\infty} l \cdot Z_l^*/Z_l < \infty$ there is some N such that $\sum_l l A_l(N)/Z_l < \gamma \cdot C^{-4} \cdot \varepsilon$. We estimate $\mu_n(x_0 > N)$. For that let $0 \leq i < n$. Then

$$\begin{aligned} S^i \nu_n(x_0 > N) &= \nu_n(x_i > N) \\ &= \sum_{j=0}^{i} \sum_{k=i}^{n-1} \nu_n(\{x|x_i > N, x[j,k] \in L_n\}) \\ &\leq \sum_{l=1}^{n} l \cdot \max_{i-l+1 \leq j \leq i} \nu_n(\{x|x_i > N, x[j,j+l) \in L_n\}). \end{aligned}$$

If there is no first return loop of length l in the set L_n, then

$$\nu_n(\{x|x_i > N, x[j,j+l) \in L_n\}) = 0,$$

and otherwise Lemma 3.8 yields

$$\nu_n(\{x|x_i > N, x[j,j+l) \in L_n\}) \leq \gamma^{-1} C^4 \frac{A_l(N)}{Z_l}.$$

This shows that

$$S^i \nu_n(x_0 > N) \leq \gamma^{-1} C^4 \cdot \sum_{l=1}^{n} l \cdot \frac{A_l(N)}{Z_l} < \varepsilon$$

and thus $\mu_n(x_0 > N) < \varepsilon$.

(b) We show that the sequence μ_n is tight. Let $\varepsilon > 0$. Fix a sequence $\varepsilon_k > 0$ such that $2 \cdot \sum_{k \geq 0} \varepsilon_k < \varepsilon$. For each $k \geq 0$ choose N_k such that $\mu_n(x_k > N_k) < \varepsilon_k$

for all n. Let $F := \{x \in X | x_k \le N_{|k|} \text{ for all } k \in \mathbb{Z}\}$. The set F is compact and $\mu_n(F^c) \le \sum_{k\in\mathbb{Z}} \mu_n(x_k > N_{|k|}) \le 2 \cdot \sum_{k\ge 0} \varepsilon_k < \varepsilon$.

(c) Since $\sum_{n=1}^{\infty} n \cdot Z_n^*/Z_n < \infty$ there is some N such that $\sum_{n>N} nZ_n^*/Z_n < \gamma \cdot C^{-4} \cdot \varepsilon$. If $n < N$ then $\mu_n\{x | i(x_s) \ne a \text{ for } 0 \le s \le N\} = 0$. Now suppose $N \le n$ and let $0 \le i < n - N$. By Lemma 3.8 we have that $\nu_n(\{x | x[j, j+l) \text{ is a first return loop at } a\}) \le \gamma^{-1} C^4 \cdot Z_l^*/Z_l$ for all $l \le n$. Thus

$$\begin{aligned} S^i \nu_n\{x | i(x_s) \ne a \text{ for } 0 \le s \le N\} &= \nu_n\{x | i(x_s) \ne a \text{ for } i \le s \le i + N\} \\ &= \sum_{j=0}^{i-1} \sum_{k=i+N}^{n-1} \nu_n(\{x | x[j,k] \text{ is a first return loop at } a\}) \\ &\le \sum_{l=N+2}^{n} l \cdot max_{0\le j<i} \cdot \nu_n(\{x | x[j, j+l) \text{ is a first return loop at } a\}) \\ &\le \gamma^{-1} C^4 \cdot \sum_{l=N+2}^{n} l \frac{Z_l^*}{Z_l} < \varepsilon \end{aligned}$$

by the choice of N. For $n - N \le i < n$ we have

$$S^i \nu_n\{x | i(x_s) \ne a \text{ for } 0 \le s \le N\} = 0. \qquad \blacksquare$$

Remark 3.11: Let $Z_0 := 1$. Since the map

$$\phi\colon (\bigcup_{k=1}^{n-1} P^*(k,a) \times P(n-k,a)) \cup P^*(n,a) \to P(n,a)$$

defined by $\phi(x,y) = z$ where $z[0,k) = x[0,k)$ and $z[k,n) = y[0,n)$ is bijective, Lemma 1.5 with $n_1 = k$ and $n_2 = n - k$ shows that

$$C^{-2} \cdot \sum_{k=1}^{n} Z_k^* \cdot Z_{n-k} \le Z_n \le C^2 \cdot \sum_{k=1}^{n} Z_k^* \cdot Z_{n-k}.$$

Thus the sequence Z_n is not necessarily a renewal sequence, but not too far from being one.

4. Equilibrium states

In the preceding section we constructed sequences of measures $\mu_n = \mu_{n,L_n}$ associated to a Z-recurrent function f and a sequence L_n of sets of first return loops. The sequence μ_n was shown to be tight whenever the sets L_n are large enough. This implies the existence of weak accumulation points [P, Thm. 6.7]. The main result of this section states the following: if the measures μ_n are good for f, then a weak accumulation point μ of the sequence μ_n is an equilibrium state if it satisfies the necessary condition $\int f^- d\mu < \infty$, i.e., if μ is a good measure for f, too.

We defined an invariant Borel probability measure μ on (X, S) to be good if $\int f^- d\mu < \infty$. There is no loss of generality in considering only good measures μ for which $h_\mu(S) + \int f d\mu$ can be evaluated and agrees with $P_{top}(f)$, just because $P_{top}(f) > -\infty$.

Definition 4.1: An equilibrium state for f is a good measure μ with $h_\mu(S) + \int f d\mu = P_{top}(f)$.

For the function $f = 0$ we have that $P_{top}(f)$ equals the Gurevic entropy of S. It is known that equilibrium states exist if $P_{top}(0) = \infty$, and for $P_{top}(0) < \infty$ an equilibrium state exists if and only if the Markov shift is positive recurrent, [G]. The existence of equilibrium states (in a slightly more general sense) was shown for Hölder continuous positive recurrent functions for which the Ruelle–Perron–Frobenius operator maps the constant function 1 to a bounded function, [S1]. We shall prove the existence of equilibrium states in a more general setting. We only assume $\sup_n D_n(f) < \infty$ instead of Hölder continuity and also cover cases where the Ruelle–Perron–Frobenius operator is of very limited use since it maps bounded functions to functions that are not real-valued (see Section 6 for a discussion on the distortion property and Hölder continuity.)

The main result of this section is the following.

THEOREM 4.2: *Suppose (X, S) is a mixing Markov shift and $f\colon X \to \mathbb{R}$ a continuous function which is Z-recurrent at vertex a, satisfies $P_{top}(f) < \infty$, and $\sup_n D_n(f) < \infty$. Let $\gamma > 0$. Assume that for each n with $P(n, a) \neq \emptyset$ a set of first return loops L_n has been chosen such that*

$$Z_{m,L_n} \geq \gamma Z_m \quad \text{for all } 1 \leq m \leq n.$$

Let

$$\nu_n := \frac{1}{Z_{n,L_n}} \cdot \sum_{x \in P(n,a,L_n)} \exp(S_n f x) \cdot \delta_x \quad \text{and} \quad \mu_n := \frac{1}{n} \sum_{i=0}^{n-1} S^i \nu_n.$$

(a) *The sequence of shift invariant probability measures μ_n is tight and has a non-empty set of weak accumulation points $\Lambda = \Lambda((\mu_n)_{n \in \mathbb{N}, P(n,a) \neq \emptyset})$.*

(b) *Suppose that all* μ_n *are good for* f *(which, for example, is satisfied if all the* L_n *are finite). Then for every* $\mu \in \Lambda$ *we have*

$$\mu \text{ is an equilibrium state for } f \text{ iff } \int f^- d\mu < \infty.$$

(c) *Suppose* $\sup_n \int f^- d\mu_n < \infty$. *Then every* $\mu \in \Lambda$ *is an equilibrium state. In particular,* f *has an equilibrium state.*

(d) *Suppose* $\inf_{x\in X} f(x) > -\infty$. *Then any sequence* μ_n *as defined above is a sequence of good measures for* f. *In particular, one might choose* L_n *to be the set of all first return loops for all* n. *Every* $\mu \in \Lambda$ *is an equilibrium state. In particular,* f *has an equilibrium state.*

Since (X, S) is mixing, $P(n, a) \neq \emptyset$ for all n large enough, thus μ_n will be defined for any large enough n. Note that the conditions $Z_{m,L_n} \geq \gamma Z_m$ for all $1 \leq m \leq n$ can be satisfied with finite sets L_n and then the associated measures μ_n are all good for f. If $\#\operatorname{Per}_n(S) < \infty$ for all n (for example, if X is locally compact), then using the whole set of first return loops at a for each L_n gives a good sequence of measures μ_n.

In Section 7 we shall use part (c) of the theorem to study the Gauss map.

Proof of Theorem 4.2: (a) The sequence μ_n is tight by Theorem 3.10, and the space is separable complete metric, thus the set of weak accumulation points is non-empty [P, Thm. 6.7].

Now suppose we have shown (b). Then

(c) If $\sup_n \int f^- d\mu_n < \infty$, then in particular all μ_n are good for f. By (a), $\Lambda \neq \emptyset$. Let $\mu \in \Lambda$. We show that μ is a good measure for f and thus an equilibrium state by (b). For each $N \in \mathbb{N}$ the function $f_N = \min(f^-, N)$ is continuous and bounded and thus $\int f_N d\mu = \lim_{k\to\infty} \int f_N d\mu_{n_k} \leq \sup_n \int f_N d\mu_n \leq \sup_n \int f^- d\mu_n$. By monotone convergence, $\int f^- d\mu = \lim_{N\to\infty} \int f_N d\mu \leq \sup \int f^- d\mu_n < \infty$.

(d) The condition implies $\sup_n \int f^- d\mu_n < \infty$ for any sequence of probability measures, thus (c) applies, in particular for the sequence where all L_n are the set of all first return loops.

(b) Let $\mu \in \Lambda$. If μ is an equilibrium state then $\int f^- d\mu < \infty$ by definition.

Now suppose $\int f^- d\mu < \infty$. Then $h_\mu(S) + \int f d\mu \leq P_{top}(f)$ by the Variational Principle, Theorem 2.4. Thus to prove that μ is an equilibrium state it remains to show that $h_\mu(S) + \int f d\mu \geq P_{top}(f)$.

Most of the rest of this section is devoted to the proof of this inequality. We will have to control certain conditional entropies. This is achieved in a preceding series of lemmata. Then, in a final step, all these estimates are put together to show the above inequality.

We start with the following useful fact; the special case $d = 1$ is the trivial part of [W1, Lemma 9.9].

LEMMA 4.3: *Let p be a discrete probability measure on a countable (possibly finite) set Y with $p(x) > 0$ for all $x \in Y$ and $g\colon Y \to \mathbb{R}$ a p-integrable function with $Z(g) := \sum_{x\in Y} e^{g(x)} < \infty$. If $d \geq 1$ is a constant such that $p(x)/p(y) \leq d \cdot e^{g(x)}/e^{g(y)}$ for all $x, y \in Y$, then*

$$|H(p) - (\log Z(g) - \int gdp)| \leq \log d.$$

Proof: By assumption $p(x)e^{g(y)} \leq de^{g(x)}p(y)$ for all x, y. Summation over y yields $p(x)/d \leq e^{g(x)}/Z(g)$ for all x. Summation over x yields $d\cdot p(y) \geq e^{g(y)}/Z(g)$ for all y. Now

$$\begin{aligned}\int gdp - \log Z(g) &= \sum_x g(x)p(x) - \sum_x (\log Z(g))p(x)\\ &= \sum_x p(x)\log\left(\frac{e^{g(x)}}{Z(g)}\right)\\ &\geq \sum_x p(x)\log\left(\frac{p(x)}{d}\right)\\ &= -H(p) - \log d.\end{aligned}$$

Similarly $\int gdp - \log Z(g) \leq -H(p) + \log d$. ∎

Let β be the zero partition of X, that is the partition elements are the sets $[i]_0 = \{x \in X | x_0 = i\}$. Let $\beta(n) = \beta \vee S^{-1}\beta \vee S^{-2}\beta \vee \cdots \vee S^{-(n-1)}\beta$ be the partition into cylinders of length n.

Now let ν_n, μ_n be defined as in Theorem 4.2 (b). Since μ_n is good, we have $h_{\mu_n}(S) + \int fd\mu_n \leq P_{top}(f) < \infty$ by the Variational Principle, Theorem 2.4. This shows that f has finite integral with respect to μ_n. Since $S^i\nu_n \leq n\cdot\mu_n$, $\int fdS^i\nu_n$ is finite for all i. Thus $1/n \cdot \int S_n f d\nu_n = \int fd\mu_n$.

LEMMA 4.4: $(1/n)\cdot H_{\nu_n}(\beta(n)) + \int fd\mu_n = (1/n)\cdot\log Z_{n,L_n}$.

Proof: Let $Y = \mathrm{supp}(\nu_n), g = S_n f, d = 1$ and $p(x) = \nu_n(x)$. Then $H_{\nu_n}(\beta(n)) = \log Z_{n,L_n} - \int S_n f d\nu_n$ by Lemma 4.3. Now use $1/n\cdot\int S_n f d\nu_n = \int fd\mu_n$. ∎

Suppose that μ is a weak accumulation point of the μ_n. To estimate $h_\mu(S)$ from below it might be impossible to use the zero-partition β, since it might have infinite entropy with respect to μ. Instead we will use some suitable finite partitions. Whatever partitions we are going to use, the next lemma gives an estimate that links to the quantity $1/n \cdot H_{\nu_n}(\beta(n))$ controlled above.

LEMMA 4.5: *Let α be some finite partition of X into closed-open sets. Let $\varepsilon > 0$. Suppose that μ is a weak limit of a subsequence μ_{n_i}. Then $h_\mu(S) \geq h_\mu(\alpha; S)$ and $h_\mu(\alpha; S) \geq (1/n)H_{\nu_n}(\beta(n)) - (1/n)H_{\nu_n}(\beta(n)|\alpha(n)) - 2\varepsilon$ for all large enough $n \in \{n_i | i \in \mathbb{N}\}$.*

Proof:

$$
\begin{aligned}
h_\mu(S) &\geq h_\mu(\alpha; S) \text{ since } \alpha \text{ is finite} \\
&\geq \frac{1}{q} H_\mu(\alpha(q)) - \varepsilon \text{ for } q \text{ large enough (supposed to be fixed from now on)} \\
&\geq \frac{1}{q} H_{\mu_n}(\alpha(q)) - 2\varepsilon \text{ by weak convergence for large enough } n = n_i \geq q \\
&\geq \frac{1}{n} H_{\mu_n}(\alpha(n)) - 2\varepsilon \text{ since } n \geq q \text{ and by [W1, Thm. 4.10]} \\
&\geq \frac{1}{n^2} \sum_{i=0}^{n-1} H_{S^i\nu_n}(\alpha(n)) - 2\varepsilon \text{ by [W1, the remark following Thm. 7.1].}
\end{aligned}
$$

Since ν_n is supported on points of period n one obtains $H_{S^i\nu_n}(\alpha(n)) = H_{\nu_n}(\alpha(n))$ for all i and $H_{\nu_n}(\beta(n) \vee \alpha(n)) = H_{\nu_n}(\beta(n))$, thus the above equals

$$
\begin{aligned}
&\frac{1}{n} H_{\nu_n}(\alpha(n)) - 2\varepsilon \\
&\qquad = \frac{1}{n} H_{\nu_n}(\alpha(n)) + \frac{1}{n} H_{\nu_n}(\beta(n)) - \frac{1}{n} H_{\nu_n}(\beta(n) \vee \alpha(n)) - 2\varepsilon \\
&\qquad = \frac{1}{n} H_{\nu_n}(\beta(n)) - \frac{1}{n} H_{\nu_n}(\beta(n)|\alpha(n)) - 2\varepsilon. \qquad \blacksquare
\end{aligned}
$$

We have to find a finite partition α that gives us sufficient control over $(1/n)H_{\nu_n}(\beta(n)|\alpha(n))$. Again, the first idea might be to use some clustering of the zero-partition β. But with such a choice we were not able to obtain the crucial entropy estimate in Lemma 4.8. However, it turns out that a partition α_L defined by a suitable finite set of return loops L works fine.

For that, suppose L is a finite set of first return loops at vertex a (not to be confused with the sequence L_n already chosen in Theorem 4.2). Let $\alpha = \alpha_L$ be the partition of X into the sets $R(w,i) := \{x \in X | x[-i, -i + |w|) = w\}$, $w \in L$, $0 \leq i < |w|$ and the "bad" set $B = X - \bigcup_{w \in L} \bigcup_{0 \leq i < |w|} R(w,i)$.

For $n \in \mathbb{N}$ let $\alpha(n) = \alpha \vee S^{-1}\alpha \vee S^{-2}\alpha \vee \cdots \vee S^{-(n-1)}\alpha$. For an atom $P \in \alpha(n)$ let $m(P)$ denote the number of times a point $x \in P$ visits the set B; thus

$m(P) = S_n 1_B(x)$ for all $x \in P$. Finally, note that the function $f^B := f \cdot 1_B$ is continuous since B is closed open.

Let $D := \sup_n D_n(f)$ and $C := \exp(D)$. Fix some constant M with $C \leq M < \infty$ and $(1/n) \log Z_n \leq M$ for all n. Such M exists since $(1/n) \log Z_n$ converges to $P_{top}(f) < \infty$ and $Z_n < \infty$ for all n.

LEMMA 4.6: *For every $P \in \alpha(n)$ with $\nu_n(P) > 0$ we have*

$$\log \left(\sum_{x \in P \cap P(n,a,L_n)} \exp(S_n f^B x) \right) \leq m(P) \cdot (D + M).$$

Proof: The statement holds trivially if $m(P) = 0$. Now assume $m(P) > 0$. Let $x \in P \cap P(n, a, L_n)$. If I is a maximal subinterval of $[0, n-1]$ such that $S^k x \in B$ for all $k \in I$, then $x|_I$ is a loop at vertex a. If $[0, n-1]$ has k such maximal subintervals, say $I_1, \ldots, I_k$, then the sum of their lengths is $m(P)$. Thus there is $y_x \in P(m(P), a)$ with $y_x[0, m(P) - 1] = x|_{I_1} \cdots x|_{I_k}$. Thus for each $1 \leq l \leq k$ we have $|\sum_{j \in I_l} S^j x - \sum_{j=a(l)}^{b(l)} S^j y_x| \leq D$ where $a(l) = 0$ if $l = 1$ and else $a(l) = |I_1| + \cdots + |I_{l-1}|$ and $b(l) = |I_1| + \cdots + |I_l| - 1$. Therefore

$$|\sum_{j \in \cup_l I_l} S^j x - \sum_{j=0}^{m(P)-1} S^j y_x| \leq k \cdot D \leq m(P) \cdot D.$$

Thus we obtain $|S_n f^B x - S_{m(P)} f y_x| \leq m(P) \cdot D$ and

$$\sum_{x \in P \cap P(n,a,L_n)} \exp(S_n f^B x) \leq \sum_{x \in P \cap P(n,a,L_n)} \exp(S_{m(P)} f y_x + m(P) \cdot D)$$
$$\leq \exp(m(P) \cdot D) Z_{m(P)}.$$

Since $\log Z_{m(P)} \leq M \cdot m(P)$ the result follows. ∎

LEMMA 4.7: *For every $P \in \alpha(n)$ with $\nu_n(P) > 0$ we have*

$$H_{\nu_n(.|P)}(\beta(n)) \leq \log \left(\sum_{x \in P \cap P(n,a,L_n)} \exp(S_n f^B x) \right) - \int S_n f^B d\nu_n(.|P) + 2m(P) \cdot D.$$

Proof: The statement is true if $m(P) = 0$. Now assume that $m(P) > 0$. Let $Y = \mathrm{supp}(\nu_n(\cdot|P)), p(x) = \nu_n(x|P)$ and $g(x) = S_n f^B x$. Then g is p-integrable, since $|S_n f^B| \leq \sum_{i=0}^{n-1} |f \circ S^i|$ and each $|f \circ S^i|$ is integrable with respect to ν_n as already observed above. Moreover, $Z(g) < \infty$ by Lemma 4.6. Now

$$\frac{p(x)}{p(y)} = \frac{\nu_n(x)}{\nu_n(y)} = \frac{\exp(S_n f x)}{\exp(S_n f y)} = \exp(S_n f x - S_n f y)$$
$$\leq \exp(S_n f^B x - S_n f^B y + 2m(P) \cdot D) = e^{2m(P)D} \frac{e^{g(x)}}{e^{g(y)}}$$

where the inequality holds, since $S^ix \in B$ iff $S^iy \in B$ and for every maximal subinterval $[i,j]$ of $[0,n-1]$ with $S^kx \notin B$ for all $i \leq k \leq j, x \in P$ we have $|\sum_{k=i}^{j} f(S^kx) - f(S^ky)| \leq D_{j-i+1}(f) \leq D$ and there are at most $m(P)+1 \leq 2m(P)$ such intervals. Thus Lemma 4.3 applies and the result follows. ∎

Combining the last two results gives the crucial estimate.

LEMMA 4.8: *For each n we have*

$$\frac{1}{n} \cdot H_{\nu_n}(\beta(n)|\alpha(n)) \leq (3D+M) \cdot \mu_n(B) - \int f^B d\mu_n.$$

Proof: By Lemma 4.6 and Lemma 4.7 for every $P \in \alpha(n)$ with $\nu_n(P) > 0$ we have

$$H_{\nu_n(.|P)}(\beta(n)) \leq m(P) \cdot (3D+M) - \int S_n f^B d\nu_n(.|P).$$

Thus

$$\begin{aligned}
\frac{1}{n} \cdot H_{\nu_n}(\beta(n)|\alpha(n)) &= \frac{1}{n} \cdot \sum_P \nu_n(P) \cdot H_{\nu_n(.|P)}(\beta(n)) \\
&\leq \frac{1}{n} \cdot \sum_P \nu_n(P) \cdot [m(P) \cdot (3D+M) - \int S_n f^B d\nu_n(.|P)] \\
&= - \int \frac{1}{n} \cdot S_n f^B d\nu_n + (3D+M) \int \frac{1}{n} \cdot S_n 1_B d\nu_n \\
&= - \int f^B d\mu_n + (3D+M)\mu_n(B).
\end{aligned}$$

This proves the lemma. ∎

Proof of Theorem 4.2(b) (continued): Recall that $\mu \in \Lambda$ is the weak limit of the subsequence μ_{n_i} and $\int f^- d\mu < \infty$. First note that f has finite μ-integral since the assumption $P_{top}(f) < \infty$ and the Variational Principle imply $\int f^+ d\mu \leq P_{top}(f) + \int f^- d\mu < \infty$.

Now let $\varepsilon > 0$. For all n large enough we have $-1/n \cdot \log \gamma < \varepsilon/2$ and, by Theorem 1.9, $1/n \cdot \log Z_n \geq P_{top}(f) - \varepsilon/2$. Thus Lemma 4.4 and the condition $Z_{n,L_n} \geq \gamma Z_n$ imply

$$\frac{1}{n} H_{\nu_n}(\beta(n)) + \int f d\mu_n \geq P_{top}(f) - \varepsilon.$$

Using Lemma 4.5 and the last inequality shows for all $n = n_i$ large enough that

$$h_\mu(S) + \int f d\mu_n + \frac{1}{n} \cdot H_{\nu_n}(\beta(n)|\alpha(n)) \geq P_{top}(f) - 3\varepsilon.$$

By Lemma 4.8, with $f^G = f - f^B$, we obtain

$$h_\mu(S) + \int f^G d\mu_{n_i} \geq P_{top}(f) - 3\varepsilon - (3D + M) \cdot \mu_{n_i}(B)$$

for all n_i large enough.

According to Theorem 3.10 there is an N with $\mu_n\{x|i(x_s) \neq a$ for $0 \leq s \leq N\} < \varepsilon^2/4$ for all n. This implies $\mu_n(\{x[j,k]$ is a first return loop for some $-N \leq j \leq 0 \leq k \leq N\}) \geq \mu_n(\{x|i(x_j) = i(x_k) = a$ for some $-N \leq j \leq 0$ and $1 \leq k \leq N+1\}) > 1 - \varepsilon^2/2$ for all n. So there is a finite set L of first return loops at a such that $\mu_n(B) \leq \varepsilon$ for all n, where B denotes the bad set of the partition $\alpha = \alpha_L$ as defined above. Since B is closed open and μ is a weak accumulation point of the μ_n, also $\mu(B) < \varepsilon$. Enlarging the set L if necessary we may assume that $\int 1_B |f| d\mu < \varepsilon$, since $\int |f| d\mu < \infty$. Since L is a finite set, B^c is a finite union of cylinder sets, thus the function $f^G = f - f^B$ is continuous and $f^G(x) = 0$ if $x \in B$. Moreover, since $D_1(f) < \infty, f^G$ is a bounded function. Thus $\int f^G d\mu_{n_i} \to \int f^G d\mu$. This convergence and $\mu_n(B) < \varepsilon$ implies

$$h_\mu(S) + \int f^G d\mu \geq P_{top}(f) - 3\varepsilon - (3D + M) \cdot \varepsilon.$$

Finally, $|\int f^B d\mu| < \varepsilon$ shows that

$$h_\mu(S) + \int f d\mu \geq P_{top}(f) - 4\varepsilon - (3D + M) \cdot \varepsilon.$$

With $\varepsilon \to 0$ this shows that $h_\mu(S) + \int f d\mu \geq P_{top}(f)$. ∎

The following is an example of a function f which satisfies all conditions of Theorem 4.2, but has no equilibrium state. Thus it can actually happen that for every tight sequence of good measures as in Theorem 4.2 the set of accumulation points of the μ_n is disjoint from the set of good measures of f. For a positive result see Section 7, where we apply Theorem 4.2(c) to the Gauss map.

Example 4.9 (inspired by [GS]): Let (X, S) be the Bernoulli shift in graph presentation with a unique vertex. Let $(a_i)_{i \in \mathbb{N}}$ be a sequence with $a_i > 0$ for all $i, \sum_i a_i = 1$ and $-\sum_i a_i \cdot \log a_i = \infty$. Define $f\colon X \to \mathbb{R}$ by $fx = \log a_i$ if $x_0 = i$. Then

$$Z_n = \sum_{i_0,\ldots,i_{n-1} \in \mathbb{N}} \exp(\log a_{i_0} + \cdots + \log a_{i_{n-1}}) = \sum_{i_0,\ldots,i_{n-1}} \prod_{k=0}^{n-1} a_{i_k} = 1.$$

Thus $P_{top}(f) = 0$ and $Z_n^* = 0$ for all $n > 1$ implies that f is Z-recurrent. We show now that f has no equilibrium state. Consider a good measure μ

for f. Let $p_i := \mu([i]_0)$ and let m denote the Bernoulli measure on X with $m([i]_0) = p_i$ for each i. Since μ is good and f is constant on each set $[i]_0$ we obtain $-\sum_i p_i \cdot \log a_i = \int f^- dm = \int f^- d\mu < \infty$. Thus m is good for f, too. For $N \in \mathbb{N}$ consider the partition α_N consisting of the sets $[i]_0, i \leq N$ and the set $\bigcup_{i>N}[i]_0$. Then $H_\mu(\alpha_N(n)) \leq n \cdot H_\mu(\alpha_N) = n \cdot H_m(\alpha_N) = H_m(\alpha_N(n))$ and thus $h_\mu(\alpha_N; S) \leq h_m(\alpha_N; S)$ for all N. Therefore $h_\mu(S) \leq h_m(S)$. Since $\int f^- dm = \int f^- d\mu$, and since m is good for f, we obtain from the Variational Principle, Theorem 2.4 that

$$h_\mu(S) + \int f d\mu \leq h_m(S) + \int f dm \leq P_{top}(f) = 0.$$

Thus $h_m(S) = -\sum_i p_i \log p_i < \infty$ and $-\sum_i p_i \log p_i + \sum_i p_i \cdot \log a_i = h_m(S) + \int f dm \leq 0 = \log(\sum_i a_i)$. Since by assumption $-\sum_i a_i \cdot \log a_i = \infty$, this implies $p_i \neq a_i$ for some i and $h_m(S) + \int f dm < 0$ ([W1, Lemma 9.9, with a_i replaced by $\log a_i$], which holds for countable probability vectors with $-\sum_i p_i \log p_i < \infty$ and $-\sum_i p_i \cdot \log a_i < \infty$). Thus $h_\mu(S) + \int f d\mu < P_{top}(f)$ and μ is not an equilibrium state for f. ∎

Remark 4.10: In certain special cases (including the Bernoulli shift) more elementary constructions of equilibrium states are possible. We would like to sketch a result, but we will not go into details, since the proofs, although more elementary, are certainly less elegant than the one given above (and we even need stronger conditions than in Theorem 4.2). The main idea is to consider, for a fixed vertex a, the following collection of subsets of X (by Theorem 3.10 it will be an almost sure partition with respect to all measures under consideration). For every first return loop w at vertex a and every $0 \leq i < |w|$ let $R(w,i) := \{x \in X | x[-i, -i + |w|) = w\}$. Let α denote the collection of all these disjoint sets $R(w,i)$. Let w^∞ denote the point $x \in \mathrm{Per}_{|w|}(S)$ with $x[0, |w|) = w$. Define a vector $p = (p_R)_{R \in \alpha}$ by

$$p_R := \frac{\exp(S_k f w^\infty)}{Z_k} \quad \text{for } R = R(w,i) \in \alpha \text{ with } k = |w|.$$

One interesting feature is that f is Z-recurrent at vertex a iff $\sum_{R \in \alpha} p_R < \infty$. This is a simple calculation:

$$\begin{aligned}\sum_{R \in \alpha} p_R &= \sum_{k=1}^{\infty} \sum_{w,|w|=k} \sum_{0 \leq i < k} p_{R(w,i)} = \sum_{k=1}^{\infty} k \cdot \sum_{w,|w|=k} \frac{\exp(S_k f w^\infty)}{Z_k} \\ &= \sum_{k=1}^{\infty} k \cdot \frac{Z_k^*}{Z_k}.\end{aligned}$$

If f satisfies the conditions of Theorem 4.2 and if additionally

$$\sup f < \infty, \quad \sup \int f^- d\mu_n < \infty \quad \text{and} \quad H(p) := -\sum_{R\in\alpha} p_R \log p_R < \infty,$$

where the μ_n are defined with L_n equal to the set of all loops (for all n), then one can show that $H_\mu(\alpha) < \infty$, thus $h_\mu(S) \geq h_\mu(\alpha; S)$. This can be used to show that any weak accumulation point of the μ_n is an equilibrium state for f. It can be shown that the above conditions are satisfied for any continuous function f with $\sup_n D_n(f) < \infty, P_{top}(f) < \infty, \sup f < \infty, \sup \int f^- d\mu_n < \infty$ defined on a mixing Markov shift (X, S) given by a graph presentation such that there is a vertex a and a number $K \in \mathbb{N}$ such that $|w| \leq K$ for all first return loops w at a.

5. One-sided Markov shifts

Let $G = (V, E)$ be a countable strongly connected directed graph. Let (X', S') be the transitive one-sided Markov shift defined by G, i.e.,

$$X' := \{x' = (x'_n) \in E^{\mathbb{N}\cup\{0\}} | t(x_n) = i(x_{n+1}) \text{ for all } n \geq 0\},$$

where t and i denote the terminal and initial vertex maps. Let f': $X' \to \mathbb{R}$ be a continuous function. Let the topological pressure of f' be defined by the same formula as in the two-sided case, see Definition 1.2. Theorem 1.9 holds in the one-sided case, too. The definition of the distortion (Definition 1.3) as well as the definition of good measures (Definition 2.1) carry over.

Let (X, S) be the transitive two-sided Markov shift given by G. Let π: $X \to X'$ be defined by $\pi(x) = (x_n)_{n\geq 0}$, where $x = (x_n)_{n\in\mathbb{Z}}$. Then π is continuous, onto, and $\pi S = S'\pi$. Let f: $X \to \mathbb{R}$ be defined by $f = f'\pi$. We call (X, S) and f the two-sided version of (X', S') and f'. Note that $D_n(f) = D_n(f')$ for all n.

THEOREM 5.1 (Variational Principle for one-sided Markov shifts): *Let (X', S') be a one-sided transitive Markov shift and let f': $X' \to \mathbb{R}$ be a function with $D_n(f')/n \to 0$. Then*

$$P_{top}(f', S') = P_{measure}(f', S').$$

Proof: Let (X, S) and f be the two-sided version of (X', S') and f'. For any vertex a the map π induces a bijection between $\mathrm{Per}_S(n, a)$ and $\mathrm{Per}_{S'}(n, a)$, thus $P_{top}(f', S') = P_{top}(f, S)$. Since $D_n(f) = D_n(f')$ for all n, by Theorem 2.4 it remains to show that $P_{measure}(f', S') = P_{measure}(f, S)$. Given a good measure μ for f on let $\mu' = \pi\mu$ denote its image under π. Then $S'\mu' = \mu'$ and obviously

$h_\mu(S) = h_{\mu'}(S')$. Now $f^- = (f')^- \circ \pi$ implies that μ' is a good measure for f' and that $\int f d\mu = \int f' d\mu'$. Thus $h_\mu(S) + \int f d\mu = h_{\mu'}(S') + \int f' d\mu'$. On the other hand, the Kolmogoroff consistency theorem shows that for any measure μ' on X' which is good for f' there is a unique S-invariant measure μ on X with $\mu(\{x \in X | x_n = a_0, \ldots, x_{n+m} = a_m\} := \mu'(\{x' \in X' | x'_0 = a_0, \ldots, x'_m = a_m\}$ for every $a_0, \ldots, a_m \in E, m \geq 0, n \in \mathbb{Z}$. Again $f^- = (f')^- \circ \pi$ implies that μ is a good measure for f. Since $\mu' = \pi\mu$ the above reasoning shows $h_\mu(S) + \int f d\mu = h_{\mu'}(S') + \int f' d\mu'$. Taking the suprema over all good measures on both sides gives $P_{measure}(f', S') = P_{measure}(f, S)$. ∎

COROLLARY 5.2: *Let (X', S') be a one-sided transitive Markov shift and let $f'\colon X' \to \mathbb{R}$ be a continuous function. Let (X, S) and f be the two-sided version of (X', S') and f'. Then f has an equilibrium state iff f' has an equilibrium state.*

Proof: This follows from the proof of Theorem 5.1. ∎

Thus, with the obvious notion of Z-recurrence Theorems 4.2 applies in the one-sided setting, too.

6. The distortion properties

This section provides estimates for the distortions $D_n(f)$ via the so-called variations $V_n(f)$ of f. In particular, the commonly used condition of Hölder continuity implies $\sup_n D_n(f) < \infty$. However, we give an example where our theorems apply but f is not Hölder continuous and not even uniformly continuous, that is $V_n(f)$ does not converge to 0. This is meant to illustrate the fact that our results apply to a wider class of functions.

Definition 6.1: The nth variation of f is

$$V_n(f) := \sup\{|fx - fy| \,|\, x[-n, n] = y[-n, n]\}.$$

Note that $V_{n+1}(f) \leq V_n(f)$ and $V_0(f) = D_1(f)$. Clearly $V_n(f) \to 0$ iff f is uniformly continuous with respect to the standard metric

$$d(x, y) = 2^{-\min\{|n| \,|\, x_n \neq y_n, n \in \mathbb{Z}\}}$$

whenever $x, y \in X, x \neq y$.

PROPOSITION 6.2: *Let (X, S) be a transitive Markov shift and let $f\colon X \to \mathbb{R}$ be a continuous function. Then*

$$\text{(a)}\quad D_n(f) \le 2 \cdot \sum_{i=0}^{n-1} V_i(f),$$

$$\text{(b)}\quad V_0(f) < \infty, V_n(f) \to 0 \Rightarrow D_n(f) < \infty \text{ for all } n \text{ and } D_n(f)/n \to 0,$$

$$\text{(c)}\quad \sum_{i=0}^{\infty} V_i(f) < \infty \Rightarrow \sup_n D_n(f) < \infty.$$

Proof: (a) Consider $x, y \in X$ with $x[0, n) = y[0, n)$. For $0 \le i < n$ let $k = k(i) = \min(i, (n-1) - i)$; then $S^i x[-k, k] = S^i y[-k, k]$ and thus $|fS^i x - fS^i y| \le V_k(f)$. Thus we have

$$|S_n f x - S_n f y| \le \sum_{i=0}^{n-1} |fS^i x - fS^i y| \le \sum_{i=0}^{n-1} V_{k(i)}(f) \le 2 \sum_{i=0}^{n-1} V_i(f).$$

Given (a) the remaining results are straightforward. ∎

A function f is Hölder continuous iff there are $M \ge 0$ and $\lambda < 1$ such that $V_n(f) \le M\lambda^n$ for all $n \ge 0$.

Observation 6.3: The above proposition shows that Hölder continuous functions satisfy $\sup_n D_n(f) < \infty$.

We give an example where our theorems apply — the function f has an equilibrium state by Theorem 4.2(d), since it is Z-recurrent and bounded from below — but f is not Hölder continuous and not even uniformly continuous, since $V_n(f) = 2$ for all n.

Example 6.4: A mixing locally compact Markov shift (X, S) and a continuous Z-recurrent function $f\colon X \to \{-1, 0, 1\}$ with $P_{top}(f) < \infty$ and $V_{n-1}(f) = D_n(f) = 2$ for all $n \ge 1$. Let G be the graph with vertex set $V = \mathbb{Z}$ and edges as follows:

- there are two edges, say c and d, from vertex 0 to vertex 1,
- there is an edge e_n from vertex n to vertex $n+1, n \ge 1$,
- there is an edge b_n from vertex n to vertex $n+1, n \le -1$,
- there is an edge a_n from vertex $(2n+1)$ to vertex $-(2n+1), n \ge 1$.

Let (X, S) be the Markov shift defined by G. Then X is locally compact and S is mixing. Define $f\colon X \to \{-1, 0, 1\}$ as follows. Let $x \in X$.

- If $x_0 \notin \{e_n | n \ge 1\}$, then let $f(x) = 0$.
- If $x_0 = e_n$, then let $f(x) = (-1)^n$ if $x_{-n} = c$ and let $f(x) = (-1)^{n+1}$ if $x_{-n} = d$.

The map is continuous, since $f(x)$ depends only on $x[-n,0]$ for some $n = n(x) \geq 0$. Since f has values in $\{-1,0,1\}, V_n(f) \leq 2$ for all n. Now let $n \geq 0$. Then consider $x, y \in X$ with $x_k = e_{n+1+k} = y_k$ for $0 \leq k \leq n, x_{-n-1} = c$ and $y_{-n-1} = d$. Then $x[-n,n] = y[-n,n]$ and $|fx - fy| = 2$. Thus $V_n(f) = 2$ for all $n \geq 0$. Now we determine $D_n(f)$. Consider $x \in X$ such that $x[0,k]$ is a first return loop at vertex 0. Then $x[1,k] = e_1 \cdots e_{2n} a_n b_{-(2n+1)} \cdots b_{-1}$ for some $n \geq 1$. Thus $S_k f(x) = \sum_{0 \leq i < k} f(S^i x) = 0$. Now consider $x, y \in X$ with $x[0,n) = y[0,n)$ and a unique $0 \leq k < n$ with $x_k \in \{c,d\}$. Then $|S_k fx|, |S_k fy| \leq 1$ and $fS^i x = fS^i y, k \leq i \leq n$, thus $|S_n fx - S_n fy| \leq 2$. Combining these two cases shows $D_n(f) \leq 2$ for all n. Let x, y with $x[0,n) = y[0,n)$ and $x_0 = y_0 = e_{2n}, x_1 = y_1 = a_n$ and $x_{-2n} = c, y_{-2n} = d$; then $|S_n fx - S_n fy| = 2$ and thus $D_n(f) = 2$.

Let a denote the vertex 0. Since for $x \in \mathrm{Per}(n,a)$ we have $S_n fx = 0$, then we get $Z_n(f,a) = \#P(n,a) < \infty$ for all n and $P_{top}(f) = h_G(S) \leq \log 2 < \infty$. Furthermore, $Z_n^*(f,a) = 2$ if $n = 4k+3, k \geq 1$, and $Z_n^*(f,a) = 0$ otherwise. Since $\#P(n,a)$ grows exponentially in n, we get that $\sum_n n Z_n^*(f,a)/Z_n(f,a) < \infty$. Thus f is Z-recurrent at a.

7. An application: the Gauss map

Although the results we present here are mainly well known, we have chosen this example to illustrate how easily our theorems apply. With $Y := (0,1) - \mathbb{Q}$ let $T\colon Y \to Y$ be the Gauss map $Tx := 1/x - [1/x]$, where $[y] := \max\{n \in \mathbb{Z} | n \leq y\}$. Then T is a countable-to-1 surjective map. For every $x \in Y$ there is a unique $k \geq 1$ such that $x \in (1/(k+1), 1/k)$, so the map T is differentiable. We will compute the topological pressure of the potential $\phi\colon Y \to \mathbb{R}$ defined by $\phi(x) = -\log|T'(x)|$ and we show that there is an equilibrium state μ of ϕ. We should mention P. Walters' work [W2] where the Gauss measure was characterized as the unique equilibrium state for $-\log|T'|$ in his sense and his method used the Ruelle–Perron–Frobenius operator. The new feature is that we characterize the Gauss measure as a limiting measure obtained from measures supported on periodic points.

The next proposition is well known, and we omit the proof.

Proposition 7.1: *There is a topological conjugacy $\tau\colon \mathbb{N}^{\mathbb{N}\cup\{0\}} \to Y$ from the one-sided Bernoulli shift $(\mathbb{N}^{\mathbb{N}\cup\{0\}}, S')$ to the dynamical system (Y,T) defined by the Gauss map. The map τ is uniformly continuous w.r.t. the standard metric on $\mathbb{N}^{\mathbb{N}\cup\{0\}}$ and the Euclidian metric on Y.*

Now let S denote the two-sided full Bernoulli shift, which means S is the left shift map on $\mathbb{N}^{\mathbb{Z}}$ and we consider this shift as given in graph presentation with a single vertex. A factor map π from S onto the one-sided shift S' is given by $\pi x = (x_0, x_1, \ldots)$. Now consider the function

$$f\colon \mathbb{N}^{\mathbb{Z}} \to \mathbb{R} \text{ given by } f = \phi\tau\pi \quad \text{where } \phi(y) = -\log|T'(y)|,\ y \in Y.$$

On a partition set $Q_k = Y \cap (1/(k+1), 1/k)$ we have $Ty = 1/y - k$, thus $T'y = -1/y^2$ and $\phi y = 2\log y$ on Q_k. Thus in particular $f \leq 0$. Elementary calculations show

CLAIM: *For each $n \geq 0$ we have $V_n(f) \leq 8 \cdot 2^{-n}$. Thus f is Hölder continuous (see Section 7 for definitions), in particular* $\sup_n D_n(f) < \infty$.

We check that $P_{top}(f) = 0$. By the above we have $C := \exp(\sup_n D_n(f)) < \infty$. Let a be the unique vertex in the graph presentation of $\mathbb{N}^{\mathbb{Z}}$.

For $a_1, \ldots, a_n \in \mathbb{N}$ let $[a_1, \ldots, a_n] := 1/(a_1 + 1/(a_2 + \cdots + 1/a_n))\cdots)$. Then $[a_1, \ldots, a_{n-1}, a_n]$ and $[a_1, \ldots, a_{n-1}, a_n + 1]$ are the two endpoints of the interval $\Delta(a_1, \ldots, a_n)$ which is the closure of $\{y \in Y | a(T^i y) = a_i, 1 \leq i < n\}$ in $[0,1]$. Let $r_n(a_1, \ldots, a_n) := \prod_{i=1}^n [a_i, a_{i+1}, \ldots, a_n]$. A simple calculation shows

$$\begin{aligned}|\Delta(a_1, \ldots, a_n)| =& |[a_1, \ldots, a_{n-1}, a_n] - [a_1, \ldots, a_{n-1}, a_n + 1]| \\ =& |[a_2, \ldots, a_{n-1}, a_n] - [a_2, \ldots, a_{n-1}, a_n + 1]| \\ & \cdot [a_1, \ldots, a_{n-1}, a_n] \cdot [a_1, \ldots, a_{n-1}, a_n + 1].\end{aligned}$$

By induction one obtains

$$|\Delta(a_1, \ldots, a_n)| = r_n(a_1, \ldots, a_{n-1}, a_n + 1) \cdot r_n(a_1, \ldots, a_{n-1}, a_n).$$

Note that by definition $\exp(S_n f z) = \prod_{i=0}^{n-1} (T^i(\tau\pi z))^2$ for all $z \in \mathbb{N}^{\mathbb{Z}}$.

Consider $z = (z_i) \in P(n, a)$; then $\tau\pi z \in \Delta(z_1, \ldots, z_n)$. Choose a sequence of points $z^k \in \mathbb{N}^{\mathbb{Z}}$ such that $\tau\pi z^k \in \Delta(z_1, \ldots, z_n)$ and $\tau\pi z^k$ converges to $[z_1, \ldots, z_n]$. Then

$$-\log C \leq S_n f z - S_n f z^k \leq \log C \Rightarrow C^{-1} \leq \frac{\exp(S_n f z)}{\prod_{i=0}^{n-1} (T^i(\tau\pi z^k))^2} \leq C \quad \text{for all } k.$$

Since $\prod_{i=0}^{n-1} (T^i(\tau\pi z^k))$ converges to $r_n(z_1, \ldots, z_n)$ we obtain

$$C^{-1} \leq \frac{\exp(S_n f z)}{r_n(z_1, \ldots, z_n)^2} \leq C.$$

Now let the z^k converge to $[z_1, \ldots, z_{n-1}, z_n + 1]$; this gives

$$C^{-1} \leq \frac{\exp(S_n f z)}{r_n(z_1, \ldots, z_{n-1}, z_n + 1)^2} \leq C.$$

The root of the product of these two estimates becomes

$$C^{-1}\cdot|\Delta(z_1,\dots,z_n)|\le \exp(S_nfz)\le C\cdot|\Delta(z_1,\dots,z_n)|.$$

Since $Z_n(f,a)=\sum_{z\in P(n,a)}\exp(S_nfz)$ and since the $|\Delta(z_1,\dots,z_n)|$ add up to 1, we obtain $C^{-1}\le Z_n(f,a)\le C$. This implies $P_{top}(f,a)=0$.

Since $Z_1^*(f,a)=Z_1(f,a)$ and $Z_n^*(f,a)=0$ for $n>1$, we get that f is Z-recurrent. Since $\nu_n:=1/Z_n\cdot\sum_{z\in P(n,a)}\exp(S_nfz)\cdot\delta_z$ and $\mu_n:=1/n\sum_{i=0}^{n-1}S^i\nu_n$, we have $\mu_n=\nu_n$ in the present case and

$$\mu_n(\{z\in X|z_0=N\})\le C^4\frac{1}{Z_1}\cdot\exp(f(N^\infty))\le\frac{M}{N^2}\quad\text{with } M:=\frac{C^4}{Z_1}.$$

By Lemma 3.8, we obtain

$$\begin{aligned}\int f^-d\mu_n&\le\sum_N\sup\{f^-(z)|z_0=N\}\cdot\mu_n(\{z\in\mathbb{N}^{\mathbb{Z}}|z_0=N\})\\&\le\sum_N\frac{M\cdot 2\log(N+1)}{N^2}.\end{aligned}$$

Thus $\sup\int f^-d\mu_n<\infty$ and, by Theorem 4.2(c), any weak accumulation point of the tight sequence μ_n is an equilibrium state for f. Now let $\rho_n=\tau\pi\mu_n$. We claim that ρ_n converges weakly to the Gauss measure m. By Theorem 4.2(c) the sequence μ_n contains a subsequence μ_{n_k} that converges weakly to a probability measure μ and μ is an equilibrium state for f. Thus, since τ,π are continuous, $\rho_{n_k}\to\tau\pi\mu$ weakly. The proof of Theorem 5.1 shows that $\pi\mu$ is an equilibrium state for $f'=\phi\tau$. Since τ is a conjugacy, $\tau\pi\mu$ is an equilibrium state for ϕ. It is well known that the Gauss measure m is the unique equilibrium state for ϕ, thus $\rho_{n_k}\to m$ weakly. The above argument applied to a subsequence μ_{k_i} of the μ_n shows that ρ_{k_i} contains a subsequence $\rho_{k_{i(j)}}$ that converges weakly to m, thus the whole sequence ρ_n converges weakly to m. Hence the Gauss measure has been obtained as a weak limit of a sequence of measures which are supported on periodic points. ∎

8. Defining pressure via all periodic points

In Definition 1.2 of topological pressure only periodic points that visit a fixed vertex a are used. Are there classes of Markov shifts where one can use all periodic points when defining pressure? This is certainly not true in general; there are simple examples of Markov shifts with finite entropy and an infinite number of fixed points, thus one cannot use all periodic points to compute $P_{top}(0)$. But the answer is positive for so-called finite-range systems; see Proposition 8.1. A natural weaker condition is the specification property, i.e., when there is some N such that for every pair of vertices a and b there is a path of length N from a to b. However, if the graph has specification and even if the function depends only on the zero-coordinate, using all periodic points might not give the topological pressure (Example 8.2).

PROPOSITION 8.1: *Suppose (X, S) is a mixing Markov shift given by a graph G with vertex set V and $f\colon X \to \mathbb{R}$ is a continuous function with $P_{top}(f) < \infty$ and $D_n(f)/n \to 0$. Let $Z(n) = Z(n, f) = \sum_{x \in \mathrm{Per}_n(S)} \exp(S_n f x)$. Suppose that V is finite. Then $P_{top}(f) = \lim(1/n) \log Z(n)$.*

Proof: Note that $Z(n) = \sum_{a \in V} Z_n(f, a)$ since $\mathrm{Per}_n(S) = \bigcup_{a \in V} P(n, a)$ by definition. We know that $P_{top}(f) = \lim(1/n) \log Z_n(f, a)$ for all $a \in V$ by Theorem 1.9. Let k be the cardinality of V. For each n there is an $a \in V$ such that $Z(n) \leq kZ_n(f, a)$. Let a_0 be some fixed vertex. Then $(1/n) \log Z_n(f, a_0) \leq (1/n) \log Z(n) \leq \max_{a \in V}(1/n) \log kZ_n(f, a)$ for all n, and the bounds converge to $P_{top}(f)$. ∎

Example 8.2: A mixing Markov shift (X, S) such that between any pair of vertices there is a path of length 2, a continuous function $f\colon X \to \mathbb{R}$ such that $V_n(f) = 0$ for all $n \geq 0$, $P_{top}(f) < \infty$, f is Z-recurrent but $\sum_{x \in \mathrm{Per}_n(S)} \exp(S_n f x) = \infty$ for all n. Thus the pressure cannot be calculated by considering all periodic points.

Let the Markov shift be given by the graph G with vertex set $V = \mathbb{N} \cup \{0\}$ and for each $k \geq 1$ there is

- an edge a_k from vertex 0 to vertex k,
- an edge b_k from vertex k to vertex k,
- an edge c_k from vertex k to vertex 0.

Let $x \in X$ and $k \geq 1$ such that $x_0 \in \{a_k, b_k, c_k\}$.

- If $x_0 = a_k$ then let $f(x) = 2 \log \frac{1}{k}$.
- If $x_0 = b_k$ or $x_0 = c_k$ then let $f(x) = \frac{1}{k} \cdot \log \frac{1}{k}$.

This defines a continuous function $f\colon X \to \mathbb{R}$ with $V_n(f) = 0$ for all $n \geq 0$. Thus, by Proposition 6.2 and Theorem 1.9, we have that

$$P_{top}(f) = \lim_{n \to \infty} (1/n) \cdot \log Z_n(f, 0).$$

Let $i(e)$ denote the initial vertex of an edge e. First note that if $x \in P^*(n,0) = \{x \in \mathrm{Per}_n(S) | i(x_0) = 0, i(x_k) \neq 0 \text{ for } 1 \leq k < n\}$ and $x_0 = a_k$, then $S_n fx = 2\log(1/k) + (n-1)\cdot(1/k)\cdot\log(1/k) \leq 2\log(1/k)$. Thus if $x \in P(n,0) = \{x \in \mathrm{Per}_n(S) | i(x_0) = 0\}$, then $x[0,n)$ can be decomposed into the first return loops at vertex 0, say $x[0,n) = x[0,i_1)x[i_1,i_2)\cdots x[i_{s-1},i_s)$ with $i_0 = 0, i_s = n$ and $1 \leq s \leq n$, and if $x_0 = a_{k_1}, \ldots, x_{i_{s-1}} = a_{k_s}$ then

$$\exp(S_n fx) \leq \exp\Big(\sum_{j=1}^{s} 2\log\frac{1}{k_j}\Big) = \prod_{j=1}^{s}\frac{1}{k_j^2}.$$

Thus

$$\begin{aligned} Z_n(f,0) &\leq \sum_{s=1}^{n} \sum_{0=i_0<i_1<\ldots<i_s=n} \sum_{(k_1,\ldots,k_s)} \prod_{j=1}^{s} \frac{1}{k_j^2} \\ &= \sum_{s=1}^{n} \sum_{0=i_0<i_1<\ldots<i_s=n} \prod_{j=1}^{s} \sum_{k=1}^{\infty} \frac{1}{k^2} \\ &\leq \sum_{s=1}^{n} \sum_{0=i_0<i_1<\ldots<i_s=n} M^n \quad \text{where } M := \sum_{k=1}^{\infty}\frac{1}{k^2} \\ &\leq M^n \cdot 2^n. \end{aligned}$$

Thus we get $P_{top}(f) \leq \log(2M) < \infty$.

To see that f is Z-recurrent first note that, for all $n \geq 2$,

$$Z_n^*(f,0) = \sum_{k=1}^{\infty} \exp\Big(2\log\frac{1}{k} + (n-1)\cdot\frac{1}{k}\cdot\log\frac{1}{k}\Big) \leq \sum_{k=1}^{\infty}\frac{1}{k^2}.$$

Thus it suffices to show that $Z_n(f,0)$ grows exponentially. By Proposition 1.8 we have $P_{top}(f) \geq P_{in}(f) \geq P(f|_Y)$ where Y is the SFT defined by the finite graph with the edges a_1, b_1, c_1. Note that $f|_Y = 0$. Thus $P(f|_Y)$ equals the topological entropy of Y. Since Y is mixing and non-trivial, clearly $P(f|_Y) > 0$ and $Z_n(f,0) \geq Z_n(f|_Y, 0)$ grows exponentially. Thus $\sum_n n \cdot Z_n^*/Z_n < \infty$ and f is Z-recurrent at vertex 0.

Finally, we show that $\sum_{x\in \mathrm{Per}_n(S)} \exp(S_n fx) = \infty$. Since $\mathrm{Per}_n(S)$ contains the fixed points x^k defined by $(x^k)_i = b_k$ for all $i \in \mathbb{Z}$, one calculates

$$\sum_{x\in\mathrm{Per}_n(S)} \exp(S_n fx) \geq \sum_{k=1}^{\infty} \exp\Big(\frac{n}{k}\cdot\log\frac{1}{k}\Big) = \infty.$$

There is no such example with $\#V = \infty$ and f bounded from below, since in this case $\#P(2N,a) = \infty$ and this trivially implies $P_{top}(f) = \lim 1/n \log Z(n) = \infty$.

References

[B] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics **470**, Springer-Verlag, Berlin, 1975.

[G] B. M. Gurevic, *Topological entropy of enumerable Markov chains*, Doklady Akademii Nauk SSSR **187** (1969); English transl.: Soviet Mathematics Doklady **10** (1969), 911–915.

[GS] B. M. Gurevic and S. V. Savchenko, *Thermodynamic formalism for countable symbolic Markov chains*, Russian Mathematical Survey **53** (1998), 245–344.

[K] B. Kitchens, *Symbolic Dynamics*, Springer, Berlin, 1998.

[P] K. R. Parthasaraty, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

[PP] Ya. Pesin and B. Pitskel, *Topological pressure and the variational principle for non-compact sets*, Functional Analysis and its Applications **18** (1984), 307–318.

[S1] O. Sarig, *Thermodynamic formalism for countable state Markov shifts*, Ergodic Theory and Dynamical Systems **19** (1999), 1565–1593.

[S2] O. Sarig, *Thermodynamic formalism for null recurrent potentials*, Israel Journal of Mathematics **121** (2001), 285–312.

[S3] O. Sarig, *Thermodynamic formalism for countable Markov shifts*, Thesis, Tel Aviv University, April 2000.

[R] D. Ruelle, *Thermodynamic formalism*, in *Encyclopedia of Mathematics and its Applications*, Vol. 5, Addison-Wesley, Reading, MA, 1978.

[W1] P. Walters, *An Introduction to Ergodic Theory*, GTM, Springer-Verlag, Berlin, 1981.

[W2] P. Walters, *Invariant measures and equilibrium states for some mappings which expand distances*, Transactions of the American Mathematical Society **236** (1978), 121–153.